Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

DATOS
TERRITORIO
MOVILIDAD
PREVENCIÓN
DESARROLLO

DECISIONES
BASADAS
EN EVIDENCIA



PREDICCIÓN
ESCENARIOS
ANÁLISIS ESPACIAL
GESTIÓN DE RIESGOS

TERRITORIOS
MÁS SEGUROS
COMUNIDADES
MÁS FUERTES

Autores:

Moreno Ponce Luis Alfonso
Solórzano Villegas Lucy Elizabeth
Carvajal Rivadeneira Daniel David
Álvarez Álvarez Martha Johnna
González Regalado Marlon Alexander
Cañarte Baque George Alfredo
Orejuela Mendoza Ivanova Claribel
Delgado Alvia Ronald Pastor
Regalado Jalca Julio Johnny

ESCANEA PARA
MÁS INFORMACIÓN

ISBN
9 789942 571366

INFRAESTRUCTURA · PERSONAS · TERRITORIOS · OPORTUNIDADES

Primera edición 2026

ISBN: 978-9942-571-36-6

2026, RUNAIKI Editora-Editorial Internacional.

https://runaiki. es/index. php/runaiki

Consejo Editorial:

Dra. Mayra Díaz Martínez, PhD.

RUNAIKI Editora-Editorial Internacional.

Hecho en Ecuador

Este texto ha sido sometido a un riguroso proceso de evaluación por pares externos.

Advertencia: Todos los derechos reservados. Queda prohibida la reproducción, registro o transmisión parcial o total de esta obra por cualquier sistema de recuperación de información existente o futuro, sin el permiso previo y por escrito del titular de los derechos correspondientes.

Escanea el QR
para acceder al libro

ISBN: 978-9942-571-36-6

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

Autores Investigadores:

Moreno Ponce Luis Alfonso
Ingeniero Civil
Magíster en Construcción de Obras Viales
Doctorado en Infraestructuras de Transporte y Territorio
Universitat Politècnica de València (UPV)
Valencia, España
luimopon@upv.edu.es
https://orcid.org/0009-0000-9026-1009

Solórzano Villegas Lucy Elizabeth
Ingeniera Civil
Magíster en Construcción de Obras Viales
Profesor de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
lucy.solorzano@unesum.edu.ec
https://orcid.org/0000-0002-9903-5304

Carvajal Rivadeneira Daniel David
Ingeniero Civil
Magíster en Administración Pública Mención Desarrollo
Profesor de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
daniel.carvajal@unesum.edu.ec
https://orcid.org/0000-0002-5288-5483

Álvarez Álvarez Martha Johnna
Ingeniero Civil
Magíster en Riego y Drenaje
Profesor de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
martha.alvarez@unesum.edu.ec
https://orcid.org/0000-0002-9879-0367

González Regalado Marlon Alexander
Ingeniero Civil
Magíster en Hidráulica con mención en Gestión de Recursos Hídricos
Profesor titular de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
marlon.gonzalez@unesum.edu.ec
https://orcid.org/0000-0003-2469-3270

Cañarte Baque George Alfredo
Arquitecto
Magíster en Tecnologías de Edificación
Profesor de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
george.canarte@unesum.edu.ec
https://orcid.org/0000-0003-3376-2557

Orejuela Mendoza Ivanova Claribel
Ingeniera Civil
Magíster en Gestión de Proyectos
Profesor titular de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
ivanova.orejuela@unesum.edu.ec
https://orcid.org/0000-0002-6094-6813

Delgado Alvia Ronald Pastor
Ingeniero Civil
Magíster en Construcción de Obras Viales
Profesor titular de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
ronald.delgado@unesum.edu.ec
https://orcid.org/0000-0002-2669-7542

Regalado Jalca Julio Johnny
Ingeniero en Computación y Redes
Magíster en Educación Informática
Profesor titular de la Universidad Estatal del Sur de Manabí
Jipijapa, Ecuador
julio.regalado@unesum.edu.ec
https://orcid.org/0000-0003-1669-8522

EDITORIAL RUNAIKI

Resumen

Este libro presenta un enfoque integrado para el estudio de la seguridad vial y la planificación de infraestructura en territorios en desarrollo, articulando fundamentos de ingeniería civil, análisis espacial, estadística y modelación predictiva. La obra examina la evolución del Sistema Seguro, la estructuración y calidad de datos viales, el uso de sistemas de información geográfica, la identificación de zonas críticas mediante densidad Kernel, la predicción temporal con modelos ARIMA y redes LSTM, la evaluación del riesgo mediante técnicas estadísticas y aprendizaje automático, así como la incorporación del drenaje vial, la accesibilidad y la priorización multicriterio de inversiones. Su propuesta metodológica combina datos observados, validación independiente, interpretación ingenieril y herramientas computacionales reproducibles, de modo que los resultados puedan traducirse en decisiones verificables sobre conservación, movilidad, prevención y desarrollo territorial. El libro incorpora aplicaciones asociadas a comunidades rurales de Manabí y aporta procedimientos que pueden fortalecer los proyectos Diagnóstico de infraestructura física para fomentar el desarrollo productivo del sitio Sancán del cantón Jipijapa y Diagnóstico de infraestructura física para fomentar el desarrollo productivo del sitio Joa del cantón Jipijapa, desarrollados por la Universidad Estatal del Sur de Manabí. Se plantea que los métodos de modelación, análisis territorial, proyección de demanda, evaluación de accesibilidad y priorización de intervenciones sean considerados en fases posteriores de ambos proyectos, con el propósito de ampliar sus capacidades de diagnóstico, seguimiento y planificación. La obra busca facilitar la integración entre investigación académica, gestión pública e ingeniería aplicada mediante procedimientos transparentes, actualizables y transferibles a territorios con características similares.

Palabras clave: seguridad vial; modelación predictiva; análisis espacial; infraestructura vial; desarrollo territorial.

Abstract

This book presents an integrated approach to road safety and infrastructure planning in developing territories by combining civil engineering, spatial analysis, statistics and predictive modelling. It examines the evolution of the Safe System approach, the structuring and quality control of road data, the use of geographic information systems, hotspot identification through Kernel Density Estimation, temporal forecasting with ARIMA models and LSTM networks, risk assessment using statistical techniques and machine learning, and the incorporation of road drainage, accessibility and multicriteria investment prioritisation. The proposed methodology combines observed data, independent validation, engineering interpretation and reproducible computational tools so that analytical outputs can be translated into verifiable decisions on maintenance, mobility, prevention and territorial development. The book also includes applications related to rural communities in Manabí and provides procedures that can strengthen the projects Diagnóstico de infraestructura física para fomentar el desarrollo productivo del sitio Sancán del cantón Jipijapa and Diagnóstico de infraestructura física para fomentar el desarrollo productivo del sitio Joa del cantón Jipijapa, developed by the Universidad Estatal del Sur de Manabí. It is proposed that the modelling methods, territorial analysis, demand forecasting, accessibility assessment and intervention prioritisation developed in this book be considered in subsequent phases of both projects to enhance their diagnostic, monitoring and planning capabilities. These contributions support evidence-based planning and the progressive improvement of territorial decision-making processes. The work also seeks to strengthen links between academic research, public management and applied engineering through transparent, updatable and transferable procedures for territories with similar characteristics and comparable infrastructure challenges.

Keywords: road safety; predictive modelling; spatial analysis; road infrastructure; territorial development.

.

ÍNDICE

RESUMEN ..... **V**

ABSTRACT ..... **VI**

PRÓLOGO ..... **XI**

1 FUNDAMENTOS TEÓRICOS DE LA SEGURIDAD VIAL Y LA MODELACIÓN PREDICTIVA **2**

1.2 DETERMINANTES ESTRUCTURALES DE LA SINIESTRALIDAD Y SU REPRESENTACIÓN CUANTITATIVA 8

1.3 ENFOQUES TRADICIONALES FRENTE A ENFOQUES BASADOS EN DATOS EN INGENIERÍA VIAL 13

1.4 FUNDAMENTOS DE LA ANALÍTICA PREDICTIVA APLICADA AL TRANSPORTE 19

1.5 TIPOLOGÍAS DE DATOS Y ESTRUCTURAS PARA MODELACIÓN 24

1.6 ESCALAS DE ANÁLISIS TERRITORIAL Y SU IMPLICACIÓN EN MODELOS PREDICTIVOS 30

1.7 RELACIÓN ENTRE INFRAESTRUCTURA, COMPORTAMIENTO HUMANO Y FORMULACIÓN DE MODELOS DE RIESGO 35

1.8 REFERENCIAS DEL CAPÍTULO 1 42

2 SISTEMAS DE INFORMACIÓN GEOGRÁFICA Y MODELACIÓN ESPACIAL DE ACCIDENTES **45**

2.1 PRINCIPIOS DEL ANÁLISIS ESPACIAL APLICADO A TRANSPORTE 45

2.2 REPRESENTACIÓN GEOESPACIAL DE SINIESTROS Y ESTRUCTURACIÓN DE DATOS GIS 50

2.3 TÉCNICAS DE GEORREFERENCIACIÓN Y LIMPIEZA DE DATOS ESPACIALES 55

2.4 MÉTODOS DE INTERPOLACIÓN ESPACIAL Y SU IMPLEMENTACIÓN COMPUTACIONAL 60

2.5 MODELACIÓN DE ESTRUCTURAS ESPACIALES MEDIANTE REGRESIÓN GEOGRÁFICA (GWR) 64

2.6 INTEGRACIÓN DE BASES DE DATOS TERRITORIALES EN ENTORNOS SIG 70

2.7 APLICACIÓN DE MODELOS ESPACIALES PREDICTIVOS EN PLANIFICACIÓN VIAL ... 74

2.8 REFERENCIAS DEL CAPÍTULO 2 81

3 MODELOS DE DENSIDAD KERNEL Y PREDICCIÓN DE ZONAS CRÍTICAS83

3.1 FUNDAMENTOS MATEMÁTICOS DEL KERNEL DENSITY..... 83
3.2 PARÁMETROS DE SUAVIZAMIENTO Y CALIBRACIÓN DEL MODELO..... 88
3.3 GENERACIÓN COMPUTACIONAL DE MAPAS DE CALOR DE ACCIDENTES..... 93
3.4 IDENTIFICACIÓN DE ZONAS CRÍTICAS MEDIANTE ANÁLISIS ESPACIAL 98
3.5 MODELACIÓN ESTADÍSTICA DE FRECUENCIA DE ACCIDENTES: POISSON Y BINOMIAL NEGATIVA 102
3.6 INTEGRACIÓN DE MODELOS KERNEL CON ENFOQUES BAYESIANOS ESPACIALES 108
3.7 PREDICCIÓN DINÁMICA DE HOTSPOTS EN REDES VIALES 113
3.8 REFERENCIAS DEL CAPÍTULO 3 120

4 MODELOS TEMPORALES ARIMA EN LA PREDICCIÓN DE SINIESTRALIDAD 122

4.1 SERIES TEMPORALES EN SEGURIDAD VIAL Y ESTRUCTURACIÓN DE DATOS 122
4.2 COMPONENTES DE TRAYECTORIA, ESTACIONALIDAD Y RUIDO EN SINIESTRALIDAD 127
4.3 FORMULACIÓN MATEMÁTICA DEL MODELO ARIMA 132
4.4 IDENTIFICACIÓN DE PARÁMETROS P, D Y Q MEDIANTE ANÁLISIS ESTADÍSTICO. 136
4.5 VALIDACIÓN, DIAGNÓSTICO Y AJUSTE DEL MODELO..... 141
4.6 IMPLEMENTACIÓN COMPUTACIONAL DE ARIMA CON DATOS DE SINIESTROS 146
4.7 EXTENSIONES DEL MODELO: SARIMA Y ARIMAX CON VARIABLES EXÓGENAS 151
4.8 REFERENCIAS DEL CAPÍTULO 4 159

5 REDES NEURONALES LSTM PARA PREDICCIÓN ESPACIO-TEMPORAL 161

5.1 FUNDAMENTOS DE REDES NEURONALES RECURRENTES EN TRANSPORTE..... 161
5.2 ARQUITECTURA LSTM Y MODELACIÓN DE DEPENDENCIAS TEMPORALES 165
5.3 PREPROCESAMIENTO DE DATOS MULTIVARIABLES ESPACIALES Y TEMPORALES 170
5.4 ENTRENAMIENTO, VALIDACIÓN Y AJUSTE DE MODELOS LSTM 175
5.5 COMPARACIÓN CUANTITATIVA ENTRE ARIMA Y LSTM 179
5.6 MODELOS HÍBRIDOS ARIMA-LSTM Y CNN-LSTM 185

5.7 Implementación de sistemas de predicción en tiempo real 189

6 AUDITORÍAS DE SEGURIDAD VIAL Y MODELOS DE EVALUACIÓN DEL RIESGO 196

6.1 Concepto y alcance de auditorías viales en ingeniería civil 196
6.2 Metodologías internacionales y estructuración de indicadores 201
6.3 Modelación del riesgo vial mediante regresión logística: logit y probit 207
6.4 Evaluación funcional y geométrica de carreteras con variables predictoras 210
6.5 Identificación de tramos críticos mediante árboles de decisión y Random Forest 215
6.6 Integración de modelos predictivos en auditorías viales 220
6.7 Diseño de intervenciones optimizadas mediante análisis multicriterio 225
6.8 Referencias del capítulo 6 232

7 DRENAJE VIAL Y MODELACIÓN HIDROLÓGICA EN LA SEGURIDAD VIAL 235

7.1 Fundamentos de hidráulica aplicada a infraestructura vial 235
7.2 Tipologías de drenaje superficial y subterráneo 240
7.3 Diseño hidráulico de cunetas y alcantarillas mediante modelos matemáticos 244
7.4 Relación entre escorrentía superficial y siniestralidad vial 249
7.5 Modelación hidrológica lluvia–escorrentía en carreteras 253
7.6 Simulación hidráulica de eventos extremos en infraestructura vial 257
7.7 Integración de modelos hidrológicos con SIG para análisis de riesgo 263
7.8 Referencias del capítulo 7 270

8 INFRAESTRUCTURA, TERRITORIO Y MODELACIÓN PREDICTIVA PARA EL DESARROLLO 272

8.1 Relación entre infraestructura vial y desarrollo productivo 272
8.2 Diagnóstico técnico de infraestructura mediante análisis de datos 277
8.3 Modelación de accesibilidad y conectividad territorial 283
8.4 Impacto de la red vial en sistemas económicos locales 289
8.5 Modelos predictivos para priorización de inversiones en infraestructura 293
8.6 Optimización de redes viales mediante modelos multicriterio 299
8.7 Lineamientos para planificación territorial basada en analítica avanzada 304
8.8 Referencias del capítulo 8 312

PRÓLOGO

La seguridad vial constituye uno de los campos de mayor responsabilidad para la ingeniería civil contemporánea, no solo por su relación directa con la movilidad y la funcionalidad de las redes de transporte, sino también por las consecuencias humanas, sociales y económicas derivadas de una infraestructura insuficientemente planificada, mantenida o evaluada. En territorios en desarrollo, donde la disponibilidad de información suele ser heterogénea y los recursos destinados a conservación y mejoramiento vial son limitados, resulta indispensable avanzar hacia metodologías capaces de integrar mediciones de campo, análisis espacial, estadística, inteligencia artificial y criterios técnicos de decisión. Bajo esta perspectiva se concibe Modelación Predictiva y Análisis Espacial para la Seguridad Vial. Infraestructura en Territorios en Desarrollo, una obra orientada a vincular los fundamentos tradicionales de la ingeniería con herramientas analíticas que permitan comprender, anticipar y gestionar los problemas asociados con el riesgo vial. El libro desarrolla de manera progresiva los fundamentos de seguridad vial, los sistemas de información geográfica, la modelación espacial de accidentes, la densidad Kernel, los modelos ARIMA, las redes neuronales LSTM, las auditorías de seguridad, el drenaje vial y la planificación de infraestructura mediante técnicas predictivas y multicriterio.

La obra ha sido estructurada con una intención académica y profesional. Cada capítulo busca que el lector comprenda primero el fundamento físico, estadístico o territorial del problema y, posteriormente, avance hacia su representación mediante modelos matemáticos, procedimientos computacionales y herramientas de análisis espacial. La incorporación de códigos en Python, ejercicios reproducibles, ecuaciones, tablas, mapas y gráficos predictivos responde a la necesidad de superar una enseñanza limitada a la exposición conceptual. El propósito es que estudiantes, docentes, investigadores y profesionales puedan reconstruir el procedimiento, evaluar sus supuestos, medir el error y relacionar los resultados con decisiones reales de ingeniería. La predicción, desde esta perspectiva, no sustituye la inspección técnica ni el juicio profesional; actúa como un instrumento para organizar

evidencia, comparar alternativas y orientar con mayor fundamento la asignación de recursos.

Un aspecto distintivo del libro es su vinculación con territorios rurales de Manabí y, particularmente, con las comunidades de Joa y Sancán, donde la infraestructura vial, la accesibilidad, el estado de los caminos y la conectividad mantienen una relación estrecha con las posibilidades de desarrollo productivo. El capítulo dedicado a infraestructura, territorio y modelación predictiva aborda precisamente el diagnóstico técnico, la accesibilidad, el impacto económico, la priorización de inversiones y la planificación territorial basada en analítica avanzada. En este sentido, la obra contribuye a fortalecer los proyectos Diagnóstico de infraestructura física para fomentar el desarrollo productivo del sitio Sancán del cantón Jipijapa y Diagnóstico de infraestructura física para fomentar el desarrollo productivo del sitio Joa del cantón Jipijapa, desarrollados por la Universidad Estatal del Sur de Manabí, al aportar herramientas que pueden incorporarse en fases posteriores para ampliar el diagnóstico, proyectar demandas, evaluar escenarios y establecer prioridades de intervención.

Más que presentar algoritmos de manera aislada, este libro propone una forma de razonar desde la ingeniería. Los datos adquieren valor cuando pueden vincularse con una carretera, una comunidad, una condición física medible y una decisión susceptible de verificación. La formación técnica requiere comprender tanto las capacidades como las limitaciones de los modelos, reconocer la incertidumbre y evitar que una salida computacional sea interpretada automáticamente como una solución constructiva. Desde esa visión, la obra aspira a servir como material de consulta, apoyo docente y referencia metodológica para quienes trabajan en seguridad vial, infraestructura y planificación territorial, especialmente en regiones donde mejorar la calidad de las decisiones puede traducirse directamente en carreteras más seguras, comunidades mejor conectadas y mayores oportunidades de desarrollo.

.

Los Autores

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO I

Fundamentos Teóricos de la Seguridad Vial y la Modelación Predictiva

AUTORES: Moreno Ponce Luis Alfonso

EDITORIAL RUNAIKI

1 Fundamentos Teóricos de la Seguridad Vial y la Modelación Predictiva

La seguridad vial contemporánea requiere integrar diseño geométrico, operación, comportamiento humano, gestión institucional y analítica de datos. El capítulo establece el lenguaje técnico que permite pasar del registro descriptivo de siniestros a modelos capaces de estimar exposición, frecuencia, severidad y riesgo. También diferencia asociación estadística, predicción y causalidad, porque cada una responde a preguntas profesionales distintas. Las revisiones del Sistema Seguro y de los modelos de seguridad basados en datos resaltan que la reducción de daños graves exige combinar infraestructura, velocidad, exposición y capacidad institucional; un modelo solo aporta valor cuando su validación corresponde con la decisión prevista (Khan & Das, 2024; Chai et al., 2024; Skaug et al., 2025).

Resultados de aprendizaje

- Explicar la transición desde enfoques reactivos hacia el Sistema Seguro y su relación con la ingeniería civil.
- Distinguir exposición, frecuencia, severidad y riesgo al definir variables para modelación.
- Seleccionar unidades de análisis y estructuras de datos compatibles con una evaluación predictiva.
- Comparar modelos simples y avanzados mediante validación fuera de muestra.
- Relacionar resultados cuantitativos con medidas físicas que puedan verificarse en la vía.

1.1 Evolución conceptual de la seguridad vial en ámbitos urbanos y rurales

Desde la ingeniería civil, el estudio de evolución conceptual de la seguridad vial en ámbitos urbanos y rurales exige algo más que ejecutar una rutina estadística. El análisis combina exposición, velocidad, energía del impacto y vulnerabilidad humana y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la

infraestructura que posteriormente deberá inspeccionarse (Washington et al., 2020; Khan & Das, 2024).



Figura 1.1. Cadena de riesgo y barreras de intervención del Sistema Seguro.
Nota. Elaboración propia con base en el enfoque de Sistema Seguro descrito por World Health Organization (2023), PIARC (2025) y Khan y Das (2024).

La representación permite observar que el riesgo se desarrolla mediante una secuencia que comienza con la exposición y puede avanzar hacia el conflicto vial, la colisión, la transferencia de energía y la lesión. Desde la ingeniería civil, esta lectura resulta especialmente útil porque permite asociar cada etapa con una intervención verificable: gestión de la demanda, control de velocidades, geometría indulgente, protección vehicular o capacidad de respuesta posterior al evento. El objetivo no consiste únicamente en impedir que ocurra un error humano, sino en evitar que ese error produzca consecuencias incompatibles con la tolerancia física de las personas.

La transferencia del análisis a redes urbanas de Jipijapa y corredores rurales de Manabí requiere aplicar el enfoque de Sistema Seguro con parámetros explícitos, datos versionados y una regla de decisión definida previamente. La prioridad consiste en definir intervenciones que reduzcan la probabilidad de un evento grave o limiten sus consecuencias. El desempeño estadístico se valora únicamente cuando mejora esa decisión. El resultado adquiere valor práctico cuando puede contrastarse con evidencia independiente, ya provenga de

inspección, normativa, medición de campo o un periodo posterior (World Health Organization, 2023; Lord & Mannering, 2010).

1.1.1 Del enfoque centrado en el accidente al Sistema Seguro

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de exposición, velocidad, energía del impacto y vulnerabilidad humana se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Skaug et al., 2025; Chai et al., 2024).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora el enfoque de Sistema Seguro y las barreras sucesivas de protección. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chai et al., 2024; Lord & Mannering, 2010).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: definir intervenciones que reduzcan la probabilidad de un evento grave o limiten sus consecuencias. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En redes urbanas de Jipijapa y corredores rurales de Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Chai et al., 2024; Washington et al., 2020).

1.1.2 Falibilidad humana y tolerancia biomecánica frente al impacto

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de exposición, velocidad, energía del impacto y vulnerabilidad humana se examinan mediante

rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Washington et al., 2020; Khan & Das, 2024).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante el enfoque de Sistema Seguro y las barreras sucesivas de protección se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Lord & Mannering, 2010; Skaug et al., 2025).

La aplicación del procedimiento cobra sentido al apoyar la decisión de definir intervenciones que reduzcan la probabilidad de un evento grave o limiten sus consecuencias. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En redes urbanas de Jipijapa y corredores rurales de Manabí, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Lord & Mannering, 2010; Washington et al., 2020).

1.1.3 Diferencias operativas entre redes urbanas y rurales

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran exposición, velocidad, energía del impacto y vulnerabilidad humana, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Lord & Mannering, 2010; Skaug et al., 2025).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora el enfoque de Sistema Seguro y las barreras sucesivas de protección. La secuencia continúa con el

cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Skaug et al., 2025; Washington et al., 2020).

La aplicación del procedimiento cobra sentido al apoyar la decisión de definir intervenciones que reduzcan la probabilidad de un evento grave o limiten sus consecuencias. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En redes urbanas de Jipijapa y corredores rurales de Manabí, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Skaug et al., 2025; World Health Organization, 2023).

Para evaluar la estabilidad del resultado, también se puede retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Washington et al., 2020; Chai et al., 2024).

1.1.4 Responsabilidad compartida y barreras sucesivas de protección

El desarrollo del apartado requiere trabajar con exposición, velocidad, energía del impacto y vulnerabilidad humana. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Lord & Mannering, 2010; Skaug et al., 2025).

La interpretación requiere una comprobación externa; la salida obtenida mediante el enfoque de Sistema Seguro y las barreras sucesivas de protección se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. La frecuencia bruta debe

interpretarse junto con la exposición, porque un tramo muy transitado puede acumular más eventos sin ser proporcionalmente más riesgoso. A la vez, la velocidad incrementa de forma cuadrática la energía del impacto. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Skaug et al., 2025; Washington et al., 2020).

Ejercicio 1.1. Evolución conceptual de la seguridad vial en ámbitos urbanos y rurales

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con TPDA, velocidad V85, usuarios vulnerables, severidad, barreras laterales. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Construya un índice de exposición y severidad; compare una lectura centrada en el error humano con una lectura de Sistema Seguro.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: lista de tres segmentos que requieren inspección de velocidad, protección lateral o cruce peatonal. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

1.2 Determinantes estructurales de la siniestralidad y su representación cuantitativa

El estudio de determinantes estructurales de la siniestralidad y su representación cuantitativa exige ordenar primero la evidencia y seleccionar después el algoritmo. La base reúne TPDA, V85, radio, pendiente, fricción, iluminación y accesos laterales y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (World Health Organization, 2023; Khan & Das, 2024).

La sección articula tasas por exposición, funciones de desempeño y regresiones de conteo con una finalidad profesional concreta: identificar factores físicos asociados con mayor frecuencia o severidad. En corredores de Jipijapa y carreteras secundarias de Manabí, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Khan & Das, 2024; Skaug et al., 2025).

1.2.1 Exposición vehicular y normalización de la frecuencia

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean TPDA, V85, radio, pendiente, fricción, iluminación y accesos laterales y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Washington et al., 2020; Khan & Das, 2024).

$$R = \frac{N}{TPDA \cdot 365 \cdot T} \times 10^6$$

En esta expresión, R es la tasa de siniestros por millón de vehículos, N representa el número de siniestros y T el periodo de observación en años. La

normalización evita comparar de forma directa tramos con niveles de exposición muy diferentes (Washington et al., 2020).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora tasas por exposición, funciones de desempeño y regresiones de conteo. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Lord & Mannering, 2010; Skaug et al., 2025).

1.2.2 Velocidad operativa, energía cinética y severidad

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean TPDA, V85, radio, pendiente, fricción, iluminación y accesos laterales y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Washington et al., 2020; Khan & Das, 2024).

$$E_k = \frac{1}{2}mv^2$$

La energía cinética E_k depende de la masa m y del cuadrado de la velocidad v. Por esta razón, incrementos relativamente pequeños de velocidad pueden aumentar de forma importante la energía que debe disiparse durante un impacto.

Para mantener la reproducibilidad, el análisis incorpora tasas por exposición, funciones de desempeño y regresiones de conteo dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad influye de forma no lineal en la severidad al aumentar cuadráticamente la

energía del impacto. El criterio de selección se define antes de observar el conjunto reservado para prueba (Lord & Mannering, 2010; Skaug et al., 2025).

La lectura ingenieril del resultado debe traducirse en una acción concreta: identificar factores físicos asociados con mayor frecuencia o severidad. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En corredores de Jipijapa y carreteras secundarias de Manabí, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Lord & Mannering, 2010; Washington et al., 2020).

Una comprobación adicional consiste en retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (World Health Organization, 2023; Khan & Das, 2024).

1.2.3 Geometría, fricción e iluminación como predictores físicos

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran TPDA, V85, radio, pendiente, fricción, iluminación y accesos laterales, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chai et al., 2024; Lord & Mannering, 2010).

La estimación se obtiene mediante tasas por exposición, funciones de desempeño y regresiones de conteo y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de

la velocidad y cambia la interpretación de la severidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Khan & Das, 2024; Washington et al., 2020).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: identificar factores físicos asociados con mayor frecuencia o severidad. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En corredores de Jipijapa y carreteras secundarias de Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Khan & Das, 2024; Chai et al., 2024).

Para ordenar predictores físicos y forma recomendada de medición, la Tabla 1.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 1-1. Predictores físicos y forma recomendada de medición

Variable	Unidad	Lectura técnica
TPDA	veh/día	Exposición al tránsito
V85	km/h	Velocidad operacional
Radio	m	Exigencia geométrica
Fricción	adimensional	Adherencia disponible
Iluminación	lux / categoría	Visibilidad nocturna

Nota. Elaboración propia a partir de variables recurrentes en modelación de seguridad vial.

1.2.4 Interacciones entre factores y funciones de desempeño de seguridad

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de TPDA, V85, radio, pendiente, fricción, iluminación y accesos laterales se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Khan & Das, 2024; Lord & Mannering, 2010).

La evaluación utiliza tasas por exposición, funciones de desempeño y regresiones de conteo, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad influye de forma no lineal en la severidad al aumentar cuadráticamente la energía del impacto. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (World Health Organization, 2023; Khan & Das, 2024).

El resultado es útil cuando permite identificar factores físicos asociados con mayor frecuencia o severidad. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En corredores de Jipijapa y carreteras secundarias de Manabí, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (World Health Organization, 2023; Chai et al., 2024).

La estabilidad del resultado se evalúa mediante cambios controlados en los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Chai et al., 2024; Lord & Mannering, 2010).

Ejercicio 1.2. Determinantes estructurales de la siniestralidad y representación cuantitativa

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con TPDA, V85, radio, pendiente, fricción y número de siniestros. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule tasas por exposición y ajuste una regresión de conteo básica; revise el signo y la magnitud de cada coeficiente.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: tabla que relacione cada factor físico con la variación esperada de la frecuencia de siniestros. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Ejemplo breve. Ejemplo de tasa: con 3 siniestros, TPDA=145 veh/día y 3 años, la tasa es 3/(145×365×3)×10^6 = 18,89 siniestros por millón de vehículos.

1.3 Enfoques tradicionales frente a enfoques basados en datos en ingeniería vial

Para trabajar con enfoques tradicionales frente a enfoques basados en datos en ingeniería vial de forma reproducible, se establece primero la unidad de análisis y la pregunta que orienta el procedimiento. Se consideran inventarios, aforos, siniestros, geometría, imágenes y costos, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Washington et al., 2020; Khan & Das, 2024).

En programas de conservación vial con información heterogénea, la comparación entre reglas de ingeniería, GLM, árboles y ensamblados debe apoyarse en parámetros explícitos, datos versionados y una regla de decisión definida previamente. El propósito es seleccionar el nivel de complejidad que realmente mejora la decisión; una métrica aislada no constituye por sí sola un

criterio de decisión. La utilidad profesional depende de que el resultado pueda verificarse mediante inspección, norma, medición independiente o datos de un periodo posterior no utilizado en el ajuste (World Health Organization, 2023; Lord & Mannering, 2010).

La disponibilidad de registros periódicos permite pasar de una lectura basada únicamente en inventarios y observaciones puntuales hacia una evaluación sustentada en series cuantitativas. La Figura 1.2 presenta los siniestros de tránsito registrados trimestralmente en Ecuador entre 2021 y 2025, y permite apreciar la variabilidad que puede quedar oculta cuando los datos se agregan únicamente por año. Esta representación constituye un ejemplo inicial de cómo una base administrativa puede convertirse en información útil para análisis estadístico y modelación posterior.

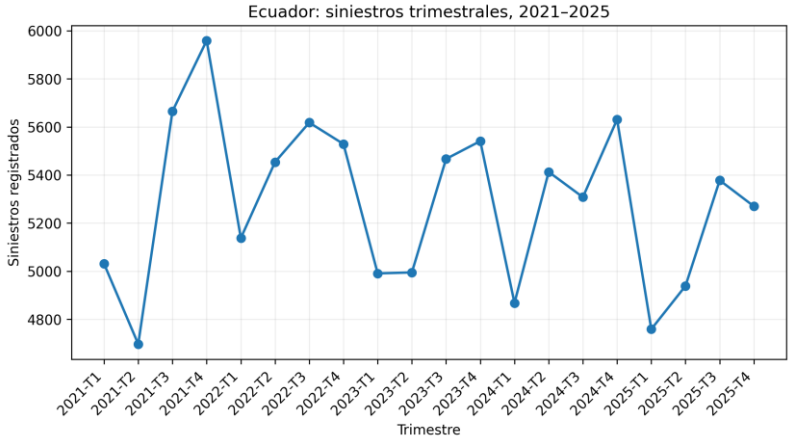


Figura 1.2. Siniestros de tránsito registrados trimestralmente en Ecuador, 2021–2025. Nota. Elaboración propia con datos del INEC provenientes de registros administrativos de la Agencia Nacional de Tránsito (ANT), 2021–2025.

La serie muestra variaciones apreciables entre trimestres, con incrementos y descensos que no siguen una progresión uniforme. Esta característica obliga a distinguir entre fluctuaciones observadas, cambios persistentes y posibles efectos asociados con exposición, movilidad o condiciones operativas. La representación gráfica constituye solamente una primera lectura; para

establecer relaciones explicativas o realizar pronósticos se requieren procedimientos estadísticos que consideren la estructura temporal de los datos.

1.3.1 Del inventario descriptivo a la evidencia cuantitativa

La evaluación se sustenta en inventarios, aforos, siniestros, geometría, imágenes y costos, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Lord & Mannering, 2010; Skaug et al., 2025).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la comparación entre reglas de ingeniería, GLM, árboles y ensamblados se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Skaug et al., 2025; Washington et al., 2020).

El resultado cuantitativo se convierte en una decisión operativa cuando permite seleccionar el nivel de complejidad que realmente mejora la decisión. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En programas de conservación vial con información heterogénea, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Skaug et al., 2025; World Health Organization, 2023).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar

una inspección o una programación preliminar (Washington et al., 2020; Chai et al., 2024).

1.3.2 Ventajas y límites de los modelos estadísticos de seguridad

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean inventarios, aforos, siniestros, geometría, imágenes y costos y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Skaug et al., 2025; Chai et al., 2024).

La interpretación requiere una comprobación externa; la salida obtenida mediante la comparación entre reglas de ingeniería, GLM, árboles y ensamblados se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. La frecuencia bruta debe interpretarse junto con la exposición, porque un tramo muy transitado puede acumular más eventos sin ser proporcionalmente más riesgoso. A la vez, la velocidad incrementa de forma cuadrática la energía del impacto. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Chai et al., 2024; Lord & Mannering, 2010).

Los resultados aportan valor técnico cuando permiten seleccionar el nivel de complejidad que realmente mejora la decisión. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En programas de conservación vial con información heterogénea, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Chai et al., 2024; Washington et al., 2020).

Para visualizar matriz conceptual de exposición y severidad para lectura preventiva del riesgo, la Figura 1.3 presenta una representación técnica que complementa la explicación y facilita su reproducción.

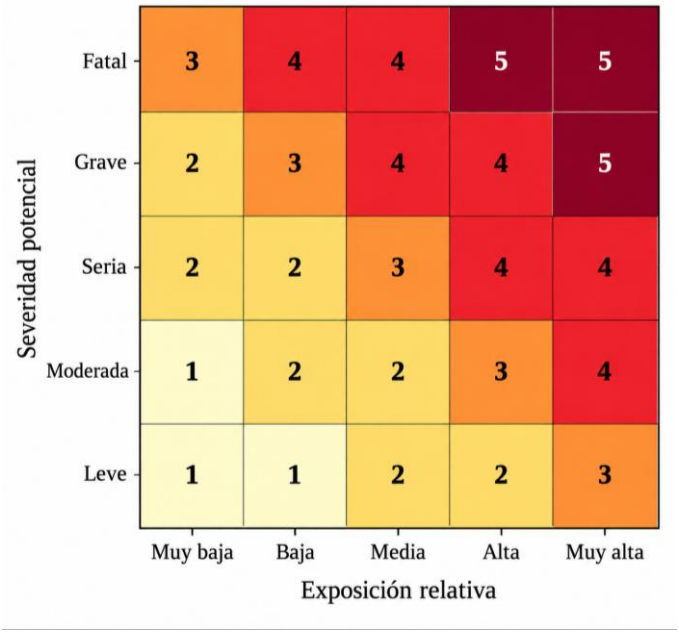


Figura 1.3. Matriz conceptual de exposición y severidad para lectura preventiva del riesgo.

Nota. Elaboración propia. Los niveles son ilustrativos y sirven para explicar la combinación entre exposición y daño potencial.

1.3.3 Aprendizaje automático frente a reglas convencionales de ingeniería

El desarrollo del apartado requiere trabajar con inventarios, aforos, siniestros, geometría, imágenes y costos. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Chai et al., 2024; Lord & Mannering, 2010).

La estimación utiliza comparación entre reglas de ingeniería, GLM, árboles y ensamblados, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba

sin utilizar información futura. La frecuencia bruta debe interpretarse junto con la exposición, porque un tramo muy transitado puede acumular más eventos sin ser proporcionalmente más riesgoso. A la vez, la velocidad incrementa de forma cuadrática la energía del impacto. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Khan & Das, 2024; Washington et al., 2020).

Los resultados aportan valor técnico cuando permiten seleccionar el nivel de complejidad que realmente mejora la decisión. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En programas de conservación vial con información heterogénea, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Khan & Das, 2024; Chai et al., 2024).

1.3.4 Trazabilidad, explicabilidad y selección del método

El desarrollo del apartado requiere trabajar con inventarios, aforos, siniestros, geometría, imágenes y costos. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (World Health Organization, 2023; Skaug et al., 2025).

El procedimiento combina la comparación entre reglas de ingeniería, GLM, árboles y ensamblados con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La frecuencia bruta debe interpretarse junto con la exposición, porque un tramo muy transitado puede acumular más eventos sin ser proporcionalmente más riesgoso. A la vez, la velocidad incrementa de forma cuadrática la energía del impacto. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Washington et al., 2020; World Health Organization, 2023).

Ejercicio 1.3. Enfoques tradicionales frente a enfoques basados en datos en ingeniería vial

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con calificación de auditoría, deficiencias de señalización, fricción, V85, siniestros. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Compare una priorización por inspección tradicional con otra obtenida mediante puntuación normalizada de datos.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: matriz de coincidencias y discrepancias entre ambos procedimientos. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

1.4 Fundamentos de la analítica predictiva aplicada al transporte

El análisis de fundamentos de la analítica predictiva aplicada al transporte, desde una perspectiva de ingeniería civil, requiere vincular el cálculo con condiciones observables de la infraestructura. El análisis combina variable objetivo, horizonte, predictores y particiones de datos y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Skaug et al., 2025; Washington et al., 2020).

La sección articula validación temporal, espacial, backtesting y métricas de error con una finalidad profesional concreta: comprobar desempeño fuera de muestra antes de priorizar. En series de siniestros y segmentos viales ecuatorianos, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Lord & Mannering, 2010; Khan & Das, 2024).

1.4.1 Definición de variable objetivo y horizonte de pronóstico

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de variable objetivo, horizonte, predictores y particiones de datos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Washington et al., 2020; Khan & Das, 2024).

La validación exige contrastar la salida obtenida mediante validación temporal, espacial, backtesting y métricas de error con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad influye de forma no lineal en la severidad al aumentar cuadráticamente la energía del impacto. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Lord & Mannering, 2010; Chai et al., 2024).

La lectura ingenieril del resultado debe traducirse en una acción concreta: comprobar desempeño fuera de muestra antes de priorizar. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En series de siniestros y segmentos viales ecuatorianos, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Lord & Mannering, 2010; Washington et al., 2020).

1.4.2 Construcción de predictores y prevención de fuga de información

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean variable objetivo, horizonte, predictores y particiones de datos y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Khan & Das, 2024; World Health Organization, 2023).

La estimación utiliza validación temporal, espacial, backtesting y métricas de error, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad influye de forma no lineal en la severidad al aumentar cuadráticamente la energía del impacto. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (World Health Organization, 2023; Chai et al., 2024).

El resultado es útil cuando permite comprobar desempeño fuera de muestra antes de priorizar. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En series de siniestros y segmentos viales ecuatorianos, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (World Health Organization, 2023; Skaug et al., 2025).

Una comprobación adicional consiste en retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Chai et al., 2024; Washington et al., 2020).

1.4.3 Particiones temporales y espaciales para evaluación fuera de muestra

La evaluación se sustenta en variable objetivo, horizonte, predictores y particiones de datos, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Chai et al., 2024; Khan & Das, 2024).

La estimación utiliza validación temporal, espacial, backtesting y métricas de error, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Khan & Das, 2024; Skaug et al., 2025).

El resultado cuantitativo se convierte en una decisión operativa cuando permite comprobar desempeño fuera de muestra antes de priorizar. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En series de siniestros y segmentos viales ecuatorianos, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Khan & Das, 2024; Lord & Mannering, 2010).

1.4.4 Métricas de error vinculadas con decisiones viales

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran variable objetivo, horizonte, predictores y particiones de datos, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta

forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (World Health Organization, 2023; Chai et al., 2024).

La estimación utiliza validación temporal, espacial, backtesting y métricas de error, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Washington et al., 2020; Skaug et al., 2025).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: comprobar desempeño fuera de muestra antes de priorizar. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series de siniestros y segmentos viales ecuatorianos, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Washington et al., 2020; World Health Organization, 2023).

Ejercicio 1.4. Fundamentos de la analítica predictiva aplicada al transporte

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con siniestros previos, TPDA, V85, lluvia, iluminación y geometría. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Separe ocho segmentos para ajuste y cuatro para prueba; estime una probabilidad de riesgo y compare con una regla simple.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: métrica de error y clasificación de los cuatro segmentos de prueba. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

1.5 Tipologías de datos y estructuras para modelación

Para analizar tipologías de datos y estructuras para modelación con rigor, primero se estructura la evidencia y luego se selecciona el procedimiento de modelación más apropiado. La base reúne marcas de tiempo, coordenadas, identificadores de segmento y atributos de vía y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (World Health Organization, 2023; Washington et al., 2020).

La transferencia a bases administrativas y levantamientos de ingeniería civil exige documentar el proceso de ETL, la normalización, los paneles y la geodatabase mediante parámetros explícitos y datos versionados. El propósito es preservar trazabilidad y compatibilidad entre fuentes; una métrica aislada no constituye por sí sola un criterio de decisión. La utilidad profesional depende de que el resultado pueda verificarse mediante inspección, norma, medición independiente o datos de un periodo posterior no utilizado en el ajuste (Khan & Das, 2024; Chai et al., 2024).

1.5.1 Datos de eventos, segmentos e intersecciones

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de marcas de tiempo, coordenadas, identificadores de segmento y atributos de vía se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma

automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Khan & Das, 2024; Chai et al., 2024).

Entre los atributos asociados a un evento pueden encontrarse variables categóricas que describen circunstancias registradas durante el siniestro, como exceso de velocidad, incumplimiento de señales u otras categorías consignadas administrativamente. Antes de incorporarlas a un modelo, estas variables deben conservar su codificación original, procedencia y periodo de observación. La Figura 1.4 presenta un ejemplo de distribución de este tipo de información para Ecuador durante 2025.

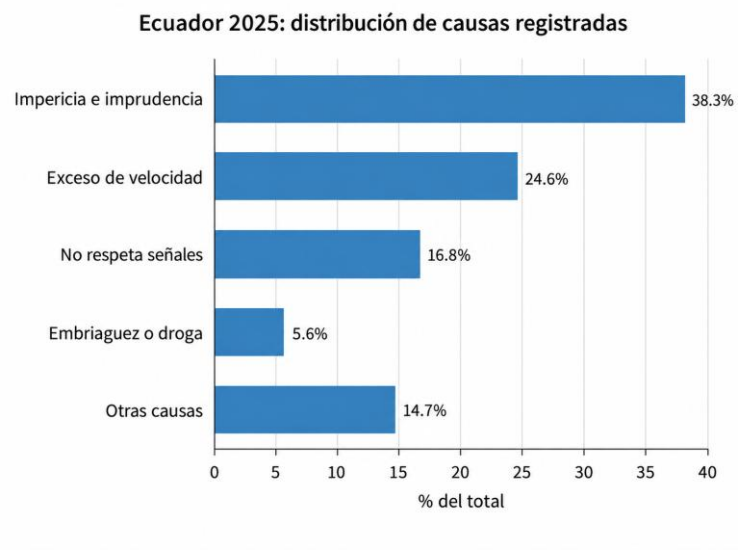


Figura 1.4. Distribución porcentual de causas registradas en Ecuador durante 2025.

Nota. Elaboración propia. Valores referenciales de presentación.

La representación muestra cómo una variable categórica de eventos puede transformarse en una distribución cuantitativa útil para la exploración inicial de una base. Sin embargo, estas categorías no deben interpretarse aisladamente como relaciones causales. Para su empleo predictivo deben vincularse con el identificador del evento, su localización, fecha, exposición vehicular y atributos físicos del segmento donde ocurrió.

La estimación se obtiene mediante ETL, normalización, paneles y geodatabases y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (World Health Organization, 2023; Chai et al., 2024).

La aplicación del procedimiento cobra sentido al apoyar la decisión de preservar trazabilidad y compatibilidad entre fuentes. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En bases administrativas y levantamientos de ingeniería civil, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (World Health Organization, 2023; Lord & Mannering, 2010).

1.5.2 Series temporales y paneles espacio-temporales

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de marcas de tiempo, coordenadas, identificadores de segmento y atributos de vía se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Skaug et al., 2025; Washington et al., 2020).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora ETL, normalización, paneles y geodatabases. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y

cambia la interpretación de la severidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chai et al., 2024; Khan & Das, 2024).

La aplicación del procedimiento cobra sentido al apoyar la decisión de preservar trazabilidad y compatibilidad entre fuentes. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En bases administrativas y levantamientos de ingeniería civil, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Chai et al., 2024; Skaug et al., 2025).

1.5.3 Información de sensores, imágenes y levantamientos de campo

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean marcas de tiempo, coordenadas, identificadores de segmento y atributos de vía y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Skaug et al., 2025; Washington et al., 2020).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante ETL, normalización, paneles y geodatabases se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Chai et al., 2024; Khan & Das, 2024).

El resultado cuantitativo se convierte en una decisión operativa cuando permite preservar trazabilidad y compatibilidad entre fuentes. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En bases administrativas y levantamientos de Ingeniería Civil, este

paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Chai et al., 2024; Skaug et al., 2025).

Para ordenar estructuras de datos y uso analítico, la Tabla 1.2 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 1-2. Estructuras de datos y uso analítico

Estructura	Unidad básica	Uso principal
Evento	Siniestro individual	Severidad y localización
Segmento	Tramo homogéneo	Frecuencia y exposición
Panel	Segmento-periodo	Modelación espacio-temporal
Serie	Periodo	Pronóstico temporal
Geodatabase	Entidad espacial	Integración territorial

Nota. Elaboración propia. La estructura debe elegirse antes de entrenar el modelo.

1.5.4 Arquitectura de bases de datos y control de versiones

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de marcas de tiempo, coordenadas, identificadores de segmento y atributos de vía se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Skaug et al., 2025; World Health Organization, 2023).

Para mantener la reproducibilidad, el análisis incorpora ETL, normalización, paneles y geodatabases dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad influye de forma no lineal en la severidad al aumentar cuadráticamente la energía del impacto. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chai et al., 2024; Skaug et al., 2025).

El resultado es útil cuando permite preservar trazabilidad y compatibilidad entre fuentes. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En bases administrativas y levantamientos de Ingeniería Civil, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Chai et al., 2024; Lord & Mannering, 2010).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Lord & Mannering, 2010; World Health Organization, 2023).

Ejercicio 1.5. Tipologías de datos y estructuras para modelación

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con fecha, identificador de segmento, coordenadas, TPDA, siniestros y estado superficial. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Organice la misma información como serie temporal, capa espacial y panel segmento-periodo; compruebe que no se dupliquen observaciones.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: tres estructuras de datos coherentes y un diccionario de variables. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.

6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

1.6 Escalas de análisis territorial y su implicación en modelos predictivos

El análisis de escalas de análisis territorial y su implicación en modelos predictivos, desde una perspectiva de ingeniería civil, requiere vincular el cálculo con condiciones observables de la infraestructura. El análisis combina segmento, intersección, corredor, zona y periodo de agregación y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Washington et al., 2020; Skaug et al., 2025).

La sección articula segmentación homogénea y análisis multiescala con una finalidad profesional concreta: evitar que la escala altere artificialmente la lectura del riesgo. En redes rurales y periurbanas con longitudes y densidades desiguales, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (World Health Organization, 2023; Khan & Das, 2024).

1.6.1 Unidad de observación y efecto de la escala espacial

En este apartado, la evidencia se organiza a partir de segmento, intersección, corredor, zona y periodo de agregación. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (World Health Organization, 2023; Lord & Mannering, 2010).

La evaluación utiliza segmentación homogénea y análisis multiescala, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad influye de forma no lineal en la severidad al aumentar cuadráticamente la energía del impacto. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Washington et al., 2020; World Health Organization, 2023).

El resultado es útil cuando permite evitar que la escala altere artificialmente la lectura del riesgo. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En redes rurales y periurbanas con longitudes y densidades desiguales, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Washington et al., 2020; Khan & Das, 2024).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Khan & Das, 2024; Skaug et al., 2025).

1.6.2 Segmentación vial y homogeneidad geométrica

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean segmento, intersección, corredor, zona y periodo de agregación y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Khan & Das, 2024; Chai et al., 2024).

La estimación utiliza segmentación homogénea y análisis multiescala, con parámetros registrados y una línea base de comparación. Antes del ajuste se

define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (World Health Organization, 2023; Chai et al., 2024).

El resultado cuantitativo se convierte en una decisión operativa cuando permite evitar que la escala altere artificialmente la lectura del riesgo. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En redes rurales y periurbanas con longitudes y densidades desiguales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (World Health Organization, 2023; Lord & Mannering, 2010).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Chai et al., 2024; World Health Organization, 2023).

1.6.3 Agregación temporal y pérdida de información

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de segmento, intersección, corredor, zona y periodo de agregación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Skaug et al., 2025; Lord & Mannering, 2010).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando segmentación homogénea y análisis multiescala como soporte analítico. Tras depurar la base, se obtienen las variables

derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. La frecuencia bruta debe interpretarse junto con la exposición, porque un tramo muy transitado puede acumular más eventos sin ser proporcionalmente más riesgoso. A la vez, la velocidad incrementa de forma cuadrática la energía del impacto. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chai et al., 2024; Washington et al., 2020).

1.6.4 Transferibilidad entre corredores y territorios

La transferibilidad entre corredores y territorios, precisa qué debe observarse y cómo debe registrarse. Para esta parte se consideran segmento, intersección, corredor, zona y periodo de agregación, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Lord & Mannering, 2010; Khan & Das, 2024).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora segmentación homogénea y análisis multiescala. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una tasa sin denominador de exposición puede sobredimensionar el riesgo de un tramo con mayor volumen. Además, la energía del impacto aumenta con el cuadrado de la velocidad y cambia la interpretación de la severidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Skaug et al., 2025; Lord & Mannering, 2010).

El resultado cuantitativo se convierte en una decisión operativa cuando permite evitar que la escala altere artificialmente la lectura del riesgo. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En redes rurales y periurbanas con longitudes y densidades desiguales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Skaug et al., 2025; Washington et al., 2020).

Ejercicio 1.6. Escalas de análisis territorial y efecto sobre los modelos

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con coordenadas, longitud de segmento, siniestros, TPDA y población expuesta. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Agregue los registros a 100 m, 500 m y 1 km; compare tasas y orden de prioridad entre escalas.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: cuadro de sensibilidad que evidencie cuánto cambia la priorización al modificar la unidad espacial. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

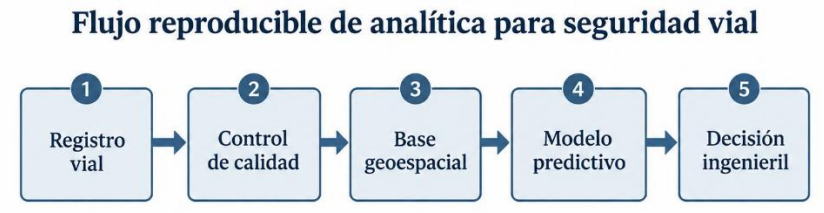


Figura 1.5. Cadena de preparación, análisis y decisión basada en datos.

Nota. Elaboración propia. La procedencia de los valores observados se indica en el párrafo de aplicación y en las referencias del capítulo.

1.7 Relación entre infraestructura, comportamiento humano y formulación de modelos de riesgo

Para analizar relación entre infraestructura, comportamiento humano y formulación de modelos de riesgo con rigor, primero se estructura la evidencia y luego se selecciona el procedimiento de modelación más apropiado. La base reúne geometría, velocidad, señalización, fricción y presencia de usuarios vulnerables y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Chai et al., 2024; Lord & Mannering, 2010).

En corredores con motociclistas, peatones y transporte mixto, la lectura conjunta de infraestructura, operación y conducta debe sustentarse en datos versionados y reglas de decisión previamente establecidas. El propósito es convertir estimaciones en medidas físicas verificables; una métrica aislada no constituye por sí sola un criterio de decisión. La utilidad profesional depende de que el resultado pueda verificarse mediante inspección, norma, medición independiente o datos de un periodo posterior no utilizado en el ajuste (Skaug et al., 2025; Khan & Das, 2024).

1.7.1 Diseño geométrico y conducta operacional del usuario

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran geometría, velocidad, señalización, fricción y presencia de usuarios vulnerables, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chai et al., 2024; Washington et al., 2020).

La evaluación utiliza lectura conjunta de infraestructura, operación y conducta, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad

influye de forma no lineal en la severidad al aumentar cuadráticamente la energía del impacto. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Khan & Das, 2024; Chai et al., 2024).

La lectura ingenieril del resultado debe traducirse en una acción concreta: convertir estimaciones en medidas físicas verificables. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En corredores con motociclistas, peatones y transporte mixto, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Khan & Das, 2024; Skaug et al., 2025).

1.7.2 Usuarios vulnerables y exposición diferencial

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran geometría, velocidad, señalización, fricción y presencia de usuarios vulnerables, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chai et al., 2024; Lord & Mannering, 2010).

La evaluación utiliza lectura conjunta de infraestructura, operación y conducta, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La exposición vehicular condiciona la lectura de la frecuencia, mientras que la velocidad influye de forma no lineal en la severidad al aumentar cuadráticamente la energía del impacto. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Khan & Das, 2024; Washington et al., 2020).

Para visualizar cadena reproducible desde la medición hasta una decisión física de ingeniería, la Figura 1.6 presenta una representación técnica que complementa la explicación y facilita su reproducción.



Figura 1.6. Cadena reproducible desde la medición hasta una decisión física de ingeniería.

Nota. Elaboración propia con orientación a procesos de modelación y validación fuera de muestra.

1.7.3 Infraestructura tolerante al error y severidad residual

Para abordar el problema se integran geometría, velocidad, señalización, fricción y presencia de usuarios vulnerables, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (World Health Organization, 2023; Chai et al., 2024).

Para sostener la interpretación, el procedimiento basado en lectura conjunta de infraestructura, operación y conducta se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. Comparar solo conteos puede inducir errores cuando los volúmenes de tránsito son distintos. La lectura de severidad también debe considerar que la energía cinética crece con el cuadrado de la velocidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Washington et al., 2020; Lord & Mannering, 2010).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: convertir estimaciones en medidas físicas verificables. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En corredores con motociclistas, peatones y transporte mixto, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa

distinción impide presentar una predicción como si fuera evidencia de campo (Washington et al., 2020; World Health Organization, 2023).

1.7.4 De la estimación del riesgo a la intervención física

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean geometría, velocidad, señalización, fricción y presencia de usuarios vulnerables y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Washington et al., 2020; World Health Organization, 2023).

La aplicación metodológica de lectura conjunta de infraestructura, operación y conducta sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. Comparar solo conteos puede inducir errores cuando los volúmenes de tránsito son distintos. La lectura de severidad también debe considerar que la energía cinética crece con el cuadrado de la velocidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Lord & Mannering, 2010; Khan & Das, 2024).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: convertir estimaciones en medidas físicas verificables. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En corredores con motociclistas, peatones y transporte mixto, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Lord & Mannering, 2010; Skaug et al., 2025).

La utilidad de un modelo de riesgo aumenta cuando el lector puede observar cómo cambia la respuesta estimada ante modificaciones de una variable física. Para ilustrar este paso, la Figura 1.7 representa una salida probabilística obtenida con una regresión logística didáctica y permite comparar condiciones superficiales bajo diferentes velocidades de operación.

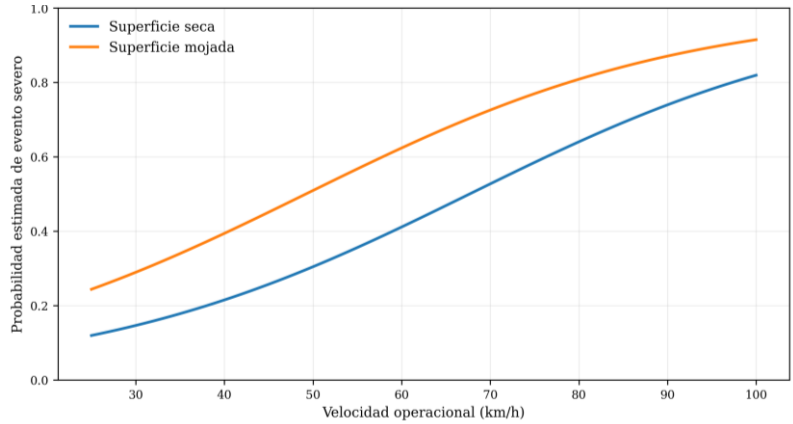


Figura 1.7. Probabilidad estimada de evento severo ante cambios de velocidad y condición superficial.

Nota. Elaboración propia a partir del Laboratorio Python 1; datos simulados con fines didácticos.

La lectura debe centrarse en la forma de la respuesta y no en los porcentajes como valores observados de una carretera real. El gráfico permite comprobar si el modelo responde de manera físicamente razonable y muestra por qué una salida predictiva debe relacionarse con velocidad, adherencia y verificación posterior en campo.

Ejercicio 1.7. Relación entre infraestructura, comportamiento humano y formulación del riesgo

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con V85, radio, ancho de carril, peatones, iluminación y severidad. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.

3) Ejecute el método. Calcule una interacción entre velocidad y radio, y otra entre exposición peatonal e iluminación; evalúe su efecto sobre un índice de riesgo.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: identificación de la combinación infraestructura-usuario que merece revisión de campo. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Practica Python 1

La siguiente practica implementa regresión logística para probabilidad de evento severo. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 1. Riesgo vial con regresión logística
# Requisitos: numpy, pandas, scikit-learn, matplotlib
import numpy as np
import pandas as pd
import matplotlib. pyplot as plt
from sklearn. model_selection import train_test_split
from sklearn. pipeline import make_pipeline
from sklearn. preprocessing import StandardScaler
from sklearn. linear_model import LogisticRegression
from sklearn. metrics import roc_auc_score, RocCurveDisplay

rng = np. random. default_rng(2026)
n = 1200

df = pd. DataFrame({
 "velocidad_kmh": rng. normal(52, 14, n). clip(20, 100),
 "tpda": rng. lognormal(np. log(500), 0.65, n). clip(50, 5000),
```

```
 "radio_m": rng. lognormal(np. log(180), 0.55, n). clip(35, 1000),
 "berma": rng. integers(0, 2, n),
 "iluminacion": rng. integers(0, 2, n),
 "superficie_mojada": rng. binomial(1, 0.22, n),
})

z = (-5.5 + 0.047 * df["velocidad_kmh"]
 + 0.00016 * df["tpda"]
 + 28 / df["radio_m"]
 - 0.55 * df["berma"]
 - 0.42 * df["iluminacion"]
 + 0.70 * df["superficie_mojada"])
p = 1 / (1 + np. exp(-z))
df["evento_severo"] = (rng. random(n) < p). astype(int)

X = df. drop(columns="evento_severo")
y = df["evento_severo"]
X_train, X_test, y_train, y_test = train_test_split(
 X, y, test_size=0.25, stratify=y, random_state=42
)

model = make_pipeline(
 StandardScaler(),
 LogisticRegression(max_iter=3000, class_weight="balanced")
)
model. fit(X_train, y_train)
prob = model. predict_proba(X_test)[:, 1]
print("AUC =", round(roc_auc_score(y_test, prob), 3))

RocCurveDisplay. from_predictions(y_test, prob)
plt. title("Curva ROC del modelo logístico")
plt. tight_layout()
plt. show()
```

Preguntas de la practica

- ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?
- ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?
- ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?
- ¿Qué información de campo agregaría para reducir incertidumbre?
- ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

1.8 Referencias del capítulo 1

World Health Organization. (2023). Global status report on road safety 2023. World Health Organization. https://www. who. int/publications/i/item/9789240086517

PIARC. (2025). Road safety manual: A manual for practitioners and decision makers on implementing safe system infrastructure. World Road Association. https://roadsafety. piarc. org/

International Transport Forum. (2024). Road safety annual report 2024. OECD Publishing.

Khan, M. N., & Das, S. (2024). Advancing traffic safety through the Safe System Approach: A systematic review. Accident Analysis & Prevention, 199, 107518. https://doi. org/10.1016/j. aap.2024.107518

Chai, A. B. Z., Lau, B. T., Tee, M. K. T., & McCarthy, C. (2024). Enhancing road safety with machine learning: Current advances and future directions in accident prediction using non-visual data. Engineering Applications of Artificial Intelligence, 137, 109086. https://doi. org/10.1016/j. engappai.2024.109086

Butt, M. S., & Shafique, M. A. (2025). A literature review: AI models for road safety for prediction of crash frequency and severity. Discover Civil Engineering, 2, 99. https://doi. org/10.1007/s44290-025-00255-3

Skaug, L., Nojoumian, M., Dang, N., & Yap, A. (2025). Road crash analysis and modeling: A systematic review of methods, data, and emerging technologies. Applied Sciences, 15(13), 7115. https://doi. org/10.3390/app15137115

American Association of State Highway and Transportation Officials. (2010). Highway safety manual (1st ed.). AASHTO.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice, 44(5), 291–305. https://doi. org/10.1016/j. tra.2010.02.001

Washington, S. P., Karlaftis, M. G., Mannering, F. L., & Anastasopoulos, P. C. (2020). Statistical and econometric methods for transportation data analysis (3rd ed.). CRC Press.

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. Accident Analysis & Prevention, 43(5), 1666–1676. https://doi. org/10.1016/j. aap.2011.03.025

Baque Parrales, E. M., Moll Martínez, R., Moreno Ponce, L. A., & Solórzano Villegas, L. E. (2022). Evaluación de la accidentalidad en la zona urbana de Jipijapa por deficiencias en la señalización vertical. Pro Sciences: Revista de Producción, Ciencias e Investigación, 6(45), 74–83. https://doi. org/10.29018/issn.2588-1000vol6iss45.2022pp74-83

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO II

Sistemas de Información Geográfica y Modelación Espacial de Accidentes

AUTOR: Solórzano Villegas Lucy Elizabeth

EDITORIAL RUNAIKI

2 Sistemas de Información Geográfica y Modelación Espacial de Accidentes

El análisis espacial transforma coordenadas aisladas en información útil para reconocer concentraciones, dependencia espacial, accesibilidad y efectos territoriales. El valor del SIG aumenta cuando la georreferenciación se combina con reglas de calidad, segmentación vial y métodos estadísticos que respetan la geometría de la red. Las revisiones recientes de aplicaciones SIG muestran que la calidad posicional, la definición de vecindad y la correspondencia con la red vial son tan importantes como el algoritmo espacial utilizado (Deressa et al., 2025; Moreno-Ponce et al., 2025).

Resultados de aprendizaje

- Estructurar información espacial de siniestros, infraestructura y exposición dentro de un SIG reproducible.
- Comprobar sistemas de referencia, precisión posicional y ajuste de eventos a la red vial.
- Aplicar autocorrelación, interpolación y regresión geográficamente ponderada con criterios de validación.
- Integrar fuentes territoriales mediante identificadores, uniones espaciales y metadatos.
- Transformar resultados cartográficos en prioridades de inspección y planificación vial.

2.1 Principios del análisis espacial aplicado a transporte

En principios del análisis espacial aplicado a transporte, la primera decisión consiste en precisar qué representa cada observación y qué pregunta deberá resolver el análisis. Se consideran coordenadas, distancias, conectividad y vecindad, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Rey et al., 2023; Oshan et al., 2019).

En redes lineales de Jipijapa y corredores interurbanos, la topología, los pesos espaciales y la autocorrelación deben analizarse con parámetros explícitos y datos cuya versión pueda ser rastreada. El análisis se orienta a definir

relaciones espaciales coherentes con la carretera, mientras que la métrica funciona como criterio de apoyo y no como un fin en sí misma. La aplicación técnica requiere una comprobación externa mediante inspección, norma, medición independiente o información posterior al periodo de ajuste (Deressa et al., 2025; Moreno-Ponce et al., 2025).

2.1.1 Topología vial, proximidad y conectividad espacial

El desarrollo del apartado requiere trabajar con coordenadas, distancias, conectividad y vecindad. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anselin, 1995; Fotheringham et al., 2002).

El procedimiento combina la topología de red, los pesos espaciales y la autocorrelación con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Oshan et al., 2019).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de definir relaciones espaciales coherentes con la carretera. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En redes lineales de Jipijapa y corredores interurbanos, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Moreno-Ponce et al., 2025; Deressa et al., 2025).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad

junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Fotheringham et al., 2002; Anselin, 1995).

2.1.2 Distancia euclidiana frente a distancia sobre la red

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de coordenadas, distancias, conectividad y vecindad se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Moreno-Ponce et al., 2025; Rey et al., 2023).

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}$$

La distancia euclidiana d_ij representa la separación directa entre dos localizaciones. En una red vial, esta medida debe compararse con la distancia recorrida sobre los enlaces, porque ambas pueden diferir de manera importante cuando existen barreras o conectividad limitada.

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante la topología de red, los pesos espaciales y la autocorrelación frente a una referencia ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Deressa et al., 2025; Moreno-Ponce et al., 2025).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: definir relaciones espaciales coherentes con la carretera. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En redes lineales de Jipijapa y corredores interurbanos, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Deressa et al., 2025; Anselin, 1995).

2.1.3 Vecindad espacial y construcción de matrices de pesos

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran coordenadas, distancias, conectividad y vecindad, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Anselin, 1995; Deressa et al., 2025).

La estimación utiliza topología de red, pesos espaciales y autocorrelación, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Moreno-Ponce et al., 2025; Anselin, 1995).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: definir relaciones espaciales coherentes con la carretera. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En redes lineales de Jipijapa y corredores interurbanos, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Moreno-Ponce et al., 2025; Fotheringham et al., 2002).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Fotheringham et al., 2002; Rey et al., 2023).

2.1.4 Autocorrelación como evidencia de dependencia territorial

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean coordenadas, distancias, conectividad y vecindad y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2025; Oshan et al., 2019).

Para mantener la reproducibilidad, el análisis incorpora la topología de red, los pesos espaciales y la autocorrelación dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de selección se define antes de observar el conjunto reservado para prueba (Deressa et al., 2025; Moreno-Ponce et al., 2025).

Los resultados deben interpretarse con una finalidad explícita: definir relaciones espaciales coherentes con la carretera. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En redes lineales de Jipijapa y corredores interurbanos, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Deressa et al., 2025; Anselin, 1995).

Ejercicio 2.1. Principios del análisis espacial aplicado al transporte

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con x, y, siniestros, TPDA, longitud y severidad. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule distancias entre segmentos, defina vecinos por proximidad y compare valores con los de sus vecinos.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: mapa o tabla de dependencia espacial y segmentos aislados. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

2.2 Representación geoespacial de siniestros y estructuración de datos GIS

La aplicación de representación geoespacial de siniestros y estructuración de datos gis requiere definir desde el inicio la unidad de análisis, el alcance temporal y la pregunta de ingeniería. Se consideran puntos de siniestros, ejes viales, segmentos y atributos, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Rey et al., 2023; Deressa et al., 2025).

La sección articula modelo vectorial y referenciación lineal con una finalidad profesional concreta: construir una unidad espacial apta para análisis y seguimiento. En eventos administrativos asociados con inventarios viales, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método

imposible de mantener, explicar o verificar en terreno (Deressa et al., 2025; Anselin, 1995).

2.2.1 Modelo de datos para puntos, líneas y segmentos viales

Para abordar el problema se integran puntos de siniestros, ejes viales, segmentos y atributos, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anselin, 1995; Moreno-Ponce et al., 2025).

Para sostener la interpretación, el procedimiento basado en modelo vectorial y referenciación lineal se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Fotheringham et al., 2002).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: construir una unidad espacial apta para análisis y seguimiento. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En eventos administrativos asociados con inventarios viales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Moreno-Ponce et al., 2025; Oshan et al., 2019).

Como verificación complementaria, se propone retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe

ocultarse detrás de una única cifra (Fotheringham et al., 2002; Deressa et al., 2025).

2.2.2 Atributos mínimos de un registro georreferenciado de siniestro

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de puntos de siniestros, ejes viales, segmentos y atributos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Rey et al., 2023; Moreno-Ponce et al., 2025).

Para mantener la reproducibilidad, el análisis incorpora modelo vectorial y referenciación lineal dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de selección se define antes de observar el conjunto reservado para prueba (Oshan et al., 2019; Fotheringham et al., 2002).

En términos de ingeniería, el análisis orienta la decisión de construir una unidad espacial apta para análisis y seguimiento. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En eventos administrativos asociados con inventarios viales, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Oshan et al., 2019; Rey et al., 2023).

2.2.3 Segmentación lineal y localización por progresiva

En este apartado, la evidencia se organiza a partir de puntos de siniestros, ejes viales, segmentos y atributos. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de

corrección aplicada a valores imposibles o incompletos (Oshan et al., 2019; Deressa et al., 2025).

La estimación utiliza modelo vectorial y referenciación lineal, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Fotheringham et al., 2002; Anselin, 1995).

Los resultados deben interpretarse con una finalidad explícita: construir una unidad espacial apta para análisis y seguimiento. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En eventos administrativos asociados con inventarios viales, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Fotheringham et al., 2002; Oshan et al., 2019).

Para ordenar campos mínimos de una base geoespacial de siniestros, la Tabla 2.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 2-1. Campos mínimos de una base geoespacial de siniestros

Campo	Tipo	Control requerido
id_evento	Texto/entero	Unicidad
fecha_hora	Fecha-hora	Consistencia temporal
x, y	Coordenadas	CRS y rango
id_segmento	Texto	Integridad referencial
severidad	Categoría	Catálogo documentado

Nota. Elaboración propia con criterios de trazabilidad geoespacial.

2.2.4 Integridad espacial entre eventos y atributos de infraestructura

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean puntos de siniestros, ejes viales, segmentos y atributos y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2025; Deressa et al., 2025).

La aplicación metodológica de modelo vectorial y referenciación lineal sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de selección se define antes de observar el conjunto reservado para prueba (Deressa et al., 2025; Anselin, 1995).

Para la ingeniería, el análisis resulta pertinente cuando permite construir una unidad espacial apta para análisis y seguimiento. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En eventos administrativos asociados con inventarios viales, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Deressa et al., 2025; Fotheringham et al., 2002).

Ejercicio 2.2. Representación geoespacial de siniestros y estructuración GIS

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con latitud, longitud, fecha, tipo de siniestro, severidad y vía. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Importe los registros en GeoPandas, asigne CRS geográfico, reproyecte a UTM 17S y realice una unión espacial con los segmentos viales.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: GeoDataFrame sin eventos fuera de la red salvo casos documentados. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

2.3 Técnicas de georreferenciación y limpieza de datos espaciales

En ingeniería civil, técnicas de georreferenciación y limpieza de datos espaciales debe abordarse como un problema técnico y no como una simple ejecución estadística. El análisis combina CRS, precisión, coordenadas y tolerancias de ajuste y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Fotheringham et al., 2002; Oshan et al., 2019).

El procedimiento adoptado emplea transformación de coordenadas, detección de anomalías y map matching y vincula cada salida con una acción verificable: evitar falsos agrupamientos causados por errores posicionales. La aplicación en registros georreferenciados sobre carreteras de Manabí obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña

con una lectura de incertidumbre y trazabilidad (Oshan et al., 2019; Deressa et al., 2025).

2.3.1 Sistemas de referencia y transformaciones de coordenadas

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran CRS, precisión, coordenadas y tolerancias de ajuste, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Oshan et al., 2019; Rey et al., 2023).

El procedimiento combina la transformación de coordenadas, la detección de anomalías y el map matching con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Fotheringham et al., 2002; Deressa et al., 2025).

La interpretación técnica orienta una acción verificable: evitar falsos agrupamientos causados por errores posicionales. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En registros georreferenciados sobre carreteras de Manabí, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Fotheringham et al., 2002; Anselin, 1995).

2.3.2 Detección de coordenadas anómalas y duplicadas

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean CRS, precisión, coordenadas y tolerancias de ajuste y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios

de codificación, precisión o cobertura entre campañas de levantamiento (Rey et al., 2023; Deressa et al., 2025).

La interpretación requiere una comprobación externa; la salida obtenida mediante la transformación de coordenadas, la detección de anomalías y el map matching se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Oshan et al., 2019; Moreno-Ponce et al., 2025).

La interpretación se orienta a una decisión técnica definida: evitar falsos agrupamientos causados por errores posicionales. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En registros georreferenciados sobre carreteras de Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Oshan et al., 2019; Fotheringham et al., 2002).

Para visualizar verificación de coordenadas y ajuste de eventos a la geometría vial, la Figura 2.1 presenta una representación técnica que complementa la explicación y facilita su reproducción.

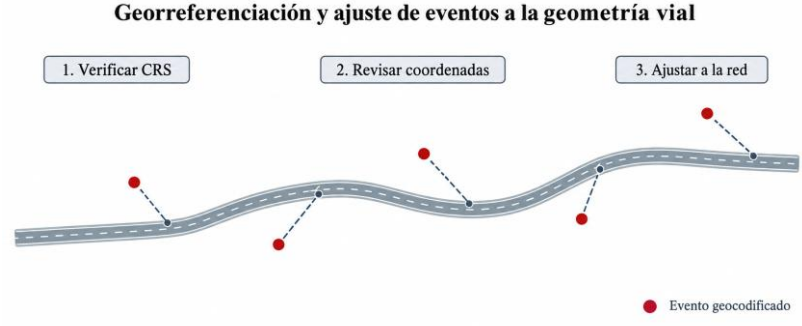


Figura 2.1. Verificación de coordenadas y ajuste de eventos a la geometría vial.

Nota. Elaboración propia. El esquema representa un procedimiento de control posicional y no ubicaciones reales de siniestros.

2.3.3 Ajuste de eventos a la red mediante map matching

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean CRS, precisión, coordenadas y tolerancias de ajuste y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Fotheringham et al., 2002; Oshan et al., 2019).

La estimación utiliza transformación de coordenadas, detección de anomalías y map matching, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Anselin, 1995; Oshan et al., 2019).

La interpretación técnica orienta una acción verificable: evitar falsos agrupamientos causados por errores posicionales. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En registros georreferenciados sobre carreteras de Manabí, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Anselin, 1995; Deressa et al., 2025).

2.3.4 Control visual y tolerancias de precisión posicional

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran CRS, precisión, coordenadas y tolerancias de ajuste, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra

la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Oshan et al., 2019; Anselin, 1995).

El procedimiento combina la transformación de coordenadas, la detección de anomalías y el map matching con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Fotheringham et al., 2002; Oshan et al., 2019).

La interpretación se orienta a una decisión técnica definida: evitar falsos agrupamientos causados por errores posicionales. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En registros georreferenciados sobre carreteras de Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Fotheringham et al., 2002; Rey et al., 2023).

Ejercicio 2.3. Georreferenciación y limpieza de datos espaciales

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con latitud, longitud, identificador, fecha, vía y punto kilométrico. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.

3) Ejecute el método. Detecte coordenadas fuera de Ecuador, duplicados exactos y eventos alejados de la vía; aplique reglas explícitas de corrección.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: bitácora que muestre valor original, regla aplicada y valor depurado. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Ejemplo breve. Ejemplo de control: una longitud -80,70° es válida para Manabí; una latitud 9851000 no lo es si la columna está declarada en grados. Antes de borrar, compruebe si corresponde a una coordenada UTM.

2.4 Métodos de interpolación espacial y su implementación computacional

Desde la ingeniería civil, el estudio de métodos de interpolación espacial y su implementación computacional exige algo más que ejecutar una rutina estadística. El análisis combina valor observado, distancia, vecindad y estructura espacial y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Fotheringham et al., 2002; Moreno-Ponce et al., 2025).

El procedimiento adoptado emplea IDW, kriging y validación cruzada y vincula cada salida con una acción verificable: estimar superficies solo cuando la variable admite continuidad espacial. La aplicación en variables ambientales o de infraestructura que acompañan al análisis vial obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (Oshan et al., 2019; Rey et al., 2023).

2.4.1 Interpolación de superficies continuas y alcance de su uso

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran valor observado, distancia, vecindad y estructura espacial, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Oshan et al., 2019; Deressa et al., 2025).

La estimación se obtiene mediante IDW, kriging y validación cruzada y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Fotheringham et al., 2002; Moreno-Ponce et al., 2025).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: estimar superficies solo cuando la variable admite continuidad espacial. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En variables ambientales o de infraestructura que acompañan al análisis vial, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Fotheringham et al., 2002; Oshan et al., 2019).

2.4.2 IDW y efecto de la distancia en la estimación

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de valor observado, distancia, vecindad y estructura espacial se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se

determina si representan un error o una condición vial real que precisamente interesa estudiar (Rey et al., 2023; Moreno-Ponce et al., 2025).

$$\hat{z}(x_0) = \frac{\sum_{i=1}^{n} d_i^{-p} z_i}{\sum_{i=1}^{n} d_i^{-p}}$$

La estimación IDW asigna mayor peso a las observaciones cercanas. El parámetro p controla la rapidez con la que disminuye esa influencia y debe evaluarse mediante validación cruzada antes de utilizar la superficie resultante.

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora IDW, kriging y validación cruzada. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de selección se define antes de observar el conjunto reservado para prueba (Oshan et al., 2019; Fotheringham et al., 2002).

La aplicación del procedimiento cobra sentido al apoyar la decisión de estimar superficies solo cuando la variable admite continuidad espacial. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En variables ambientales o de infraestructura que acompañan al análisis vial, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Oshan et al., 2019; Rey et al., 2023).

2.4.3 Kriging, estructura espacial y semivariograma

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean valor observado, distancia, vecindad y estructura espacial y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Fotheringham et al., 2002; Anselin, 1995).

La estimación utiliza IDW, kriging y validación cruzada, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Anselin, 1995; Oshan et al., 2019).

La utilidad del análisis se hace evidente cuando permite estimar superficies solo cuando la variable admite continuidad espacial. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En variables ambientales o de infraestructura que acompañan al análisis vial, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Anselin, 1995; Rey et al., 2023).

2.4.4 Validación cruzada de superficies interpoladas

La evaluación se sustenta en valor observado, distancia, vecindad y estructura espacial, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anselin, 1995; Oshan et al., 2019).

La estimación se obtiene mediante IDW, kriging y validación cruzada y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial.

Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Deressa et al., 2025).

Ejercicio 2.4. Interpolación espacial e implementación computacional

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con coordenadas y una variable continúa medida en puntos, como fricción o precipitación. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Aplique interpolación IDW con dos potencias distintas y valide retirando un punto cada vez.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: superficie interpolada y error medio de validación cruzada. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

2.5 Modelación de estructuras espaciales mediante regresión geográfica (GWR)

La ingeniería civil aborda modelación de estructuras espaciales mediante regresión geográfica (gwr) mediante la integración de evidencia física y análisis cuantitativo, no mediante una rutina estadística aislada. El análisis combina coeficientes locales, coordenadas, ancho de banda y residuos y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa

con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Rey et al., 2023; Deressa et al., 2025).

La aplicación práctica de GWR y MGWR permiten disponer la información de manera que sea posible reconocer relaciones que cambian a lo largo del territorio. El análisis se plantea para corredores donde velocidad, accesos y geometría no actúan de forma uniforme, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Deressa et al., 2025; Anselin, 1995).

2.5.1 Coeficientes locales y heterogeneidad espacial

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean coeficientes locales, coordenadas, ancho de banda y residuos y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2025; Anselin, 1995).

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k\,(u_i, v_i) x_{ik} + \varepsilon_i$$

En GWR, los coeficientes cambian con la localización (u_i, v_i). Esta formulación permite representar relaciones espaciales no constantes, aunque su interpretación requiere comprobar colinealidad local, ancho de banda y estabilidad de los parámetros.

La aplicación metodológica de GWR y MGWR sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de

selección se define antes de observar el conjunto reservado para prueba (Deressa et al., 2025; Fotheringham et al., 2002).

2.5.2 Selección del ancho de banda en GWR

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de coeficientes locales, coordenadas, ancho de banda y residuos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Fotheringham et al., 2002; Deressa et al., 2025).

El procedimiento combina GWR y MGWR con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Anselin, 1995; Fotheringham et al., 2002).

Los resultados aportan valor técnico cuando permiten reconocer relaciones que cambian a lo largo del territorio. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En corredores donde velocidad, accesos y geometría no actúan de forma uniforme, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Anselin, 1995; Oshan et al., 2019).

Para visualizar variación espacial ilustrativa de coeficientes locales en una gwr, la Figura 2.2 presenta una representación técnica que complementa la explicación y facilita su reproducción.

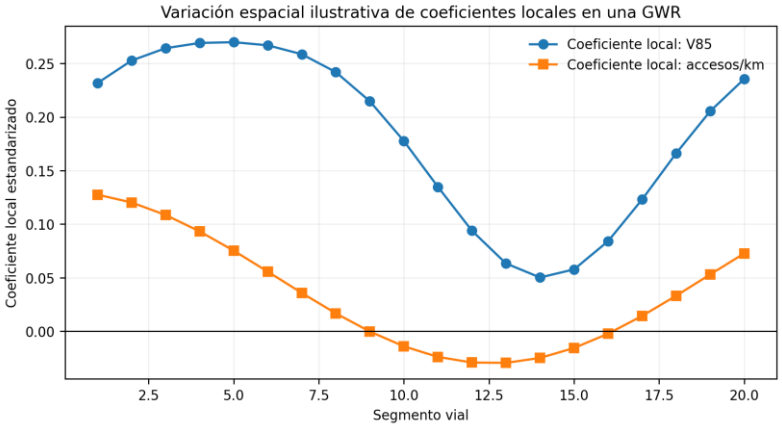


Figura 2.2. Variación espacial ilustrativa de coeficientes locales en una GWR.

Nota. Elaboración propia. Los valores son didácticos y muestran por qué un coeficiente global puede ocultar variación espacial.

2.5.3 Colinealidad local y estabilidad de parámetros

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran coeficientes locales, coordenadas, ancho de banda y residuos, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Deressa et al., 2025; Moreno-Ponce et al., 2025).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando GWR y MGWR como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de selección se define antes de observar el conjunto reservado para prueba (Rey et al., 2023; Anselin, 1995).

La interpretación técnica orienta una acción verificable: reconocer relaciones que cambian a lo largo del territorio. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En corredores donde velocidad, accesos y geometría no actúan de forma uniforme, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Rey et al., 2023; Oshan et al., 2019).

La sensibilidad del resultado se estudia mediante variaciones controladas de los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Moreno-Ponce et al., 2025; Deressa et al., 2025).

2.5.4 Comparación entre regresión global, GWR y MGWR

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran coeficientes locales, coordenadas, ancho de banda y residuos, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Oshan et al., 2019; Anselin, 1995).

Para sostener la interpretación, el procedimiento basado en GWR y MGWR se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Fotheringham et al., 2002; Oshan et al., 2019).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: reconocer relaciones que cambian a lo largo del territorio. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En corredores donde velocidad, accesos y geometría no actúan de forma uniforme, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Fotheringham et al., 2002; Rey et al., 2023).

Ejercicio 2.5. Regresión geográficamente ponderada (GWR)

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con x, y, tasa de siniestros, V85, accesos por km y radio de curva. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Estime primero una regresión global y luego regresiones locales ponderadas por distancia; compare RMSE y variación de coeficientes.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: mapa de coeficientes locales y advertencia sobre colinealidad o tamaño muestral. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Ejemplo breve. Ejemplo de lectura: si el coeficiente local de accesos/km aumenta hacia un sector, la asociación es más intensa allí; antes de proponer una obra revise colinealidad, tamaño local efectivo y residuos. El Anexo técnico B incluye un código GWR funcional sin paquetes especializados.

2.6 Integración de bases de datos territoriales en entornos SIG

Antes de modelar integración de bases de datos territoriales en entornos sig, es necesario organizar la evidencia y comprobar que responda a la pregunta técnica planteada. La base reúne identificadores, fechas, geometrías y catálogos de atributos y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Oshan et al., 2019; Rey et al., 2023).

La sección articula uniones espaciales, temporales y control de linaje con una finalidad profesional concreta: mantener una base integrada que pueda actualizarse sin duplicaciones. En SIG institucional y levantamientos universitarios, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Rey et al., 2023; Moreno-Ponce et al., 2025).

2.6.1 Identificadores persistentes para integrar fuentes heterogéneas

El desarrollo del apartado requiere trabajar con identificadores, fechas, geometrías y catálogos de atributos. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anselin, 1995; Fotheringham et al., 2002).

El procedimiento combina uniones espaciales, temporales y control de linaje con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de

parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Oshan et al., 2019).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de mantener una base integrada que pueda actualizarse sin duplicaciones. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En SIG institucional y levantamientos universitarios, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Moreno-Ponce et al., 2025; Deressa et al., 2025).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Fotheringham et al., 2002; Anselin, 1995).

2.6.2 Uniones espaciales y temporales sin duplicar observaciones

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de identificadores, fechas, geometrías y catálogos de atributos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Fotheringham et al., 2002; Oshan et al., 2019).

Para sostener la interpretación, el procedimiento basado en uniones espaciales, temporales y control de linaje se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Anselin, 1995; Oshan et al., 2019).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de mantener una base integrada que pueda actualizarse sin duplicaciones. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En SIG institucional y levantamientos universitarios, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Anselin, 1995; Deressa et al., 2025).

2.6.3 Metadatos, linaje del dato y reglas de actualización

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de identificadores, fechas, geometrías y catálogos de atributos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Fotheringham et al., 2002; Anselin, 1995).

Para sostener la interpretación, el procedimiento basado en uniones espaciales, temporales y control de linaje se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Anselin, 1995; Oshan et al., 2019).

Para ordenar estrategias de integración territorial en sig, la Tabla 2.2 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 2-2. Estrategias de integración territorial en SIG

Operación	Riesgo frecuente	Verificación
Join por ID	IDs no persistentes	Conteo antes/después

Spatial join	Duplicación por intersección	Regla de cardinalidad
Join temporal	Fechas desalineadas	Ventana documentada
Map matching	Asignación al tramo incorrecto	Distancia residual
Actualización	Sobrescritura de históricos	Versionado

Nota. Elaboración propia. Las uniones deben preservar el número y la identidad de observaciones.

2.6.4 Diseño de una geodatabase vial reproducible

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de identificadores, fechas, geometrías y catálogos de atributos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Rey et al., 2023; Moreno-Ponce et al., 2025).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora uniones espaciales, temporales y control de linaje. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de selección se define antes de observar el conjunto reservado para prueba (Oshan et al., 2019; Anselin, 1995).

El resultado cuantitativo se convierte en una decisión operativa cuando permite mantener una base integrada que pueda actualizarse sin duplicaciones. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En SIG institucional y levantamientos universitarios, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Oshan et al., 2019; Rey et al., 2023).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Deressa et al., 2025; Moreno-Ponce et al., 2025).

Ejercicio 2.6. Integración de bases territoriales en entornos SIG

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con ID de segmento, geometría, TPDA, población, drenaje, señalización y siniestros. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Normalice identificadores, realice uniones uno-a-uno y uno-a-muchos, y compruebe pérdidas antes y después de cada unión.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: geodatabase integrada con reporte de registros no emparejados. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

2.7 Aplicación de modelos espaciales predictivos en planificación vial

La comprensión de aplicación de modelos espaciales predictivos en planificación vial requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran superficies de riesgo, exposición,

infraestructura y costo; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Oshan et al., 2019; Rey et al., 2023).

La sección articula escenarios espaciales y priorización cartográfica con una finalidad profesional concreta: programar inspecciones y obras con evidencia territorial. En tramos de Chone–Flavio Alfaro y redes locales de Manabí, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Rey et al., 2023; Moreno-Ponce et al., 2025).

2.7.1 Superficies de riesgo para priorización de inspecciones

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean superficies de riesgo, exposición, infraestructura y costo y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2025; Anselin, 1995).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora los escenarios espaciales y la priorización cartográfica. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. El criterio de selección se define antes de observar el conjunto reservado para prueba (Deressa et al., 2025; Fotheringham et al., 2002).

Una aplicación espacial no queda completa con la visualización de puntos o coeficientes. La salida que interesa a la planificación es una superficie o perfil de riesgo esperado que permita localizar sectores donde conviene intensificar la inspección. La Figura 2.3 muestra este tipo de producto para un corredor didáctico.

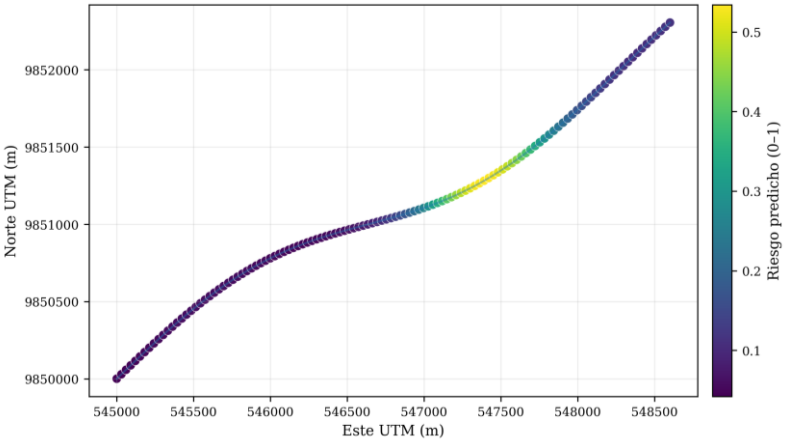


Figura 2.3. Riesgo espacial predicho a lo largo de un corredor vial.

Nota. Elaboración propia con datos sintéticos; la salida ilustra una predicción espacial y no ubicaciones reales de siniestros.

El perfil distingue sectores con mayor riesgo relativo y facilita comparar la continuidad espacial de la señal. Antes de priorizar una intervención, los máximos del modelo deben contrastarse con geometría, velocidad, accesos, condición superficial y observaciones de terreno.

2.7.2 Escenarios espaciales de intervención y reasignación del riesgo

La evaluación se sustenta en superficies de riesgo, exposición, infraestructura y costo, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anselin, 1995; Moreno-Ponce et al., 2025).

La estimación se obtiene mediante los escenarios espaciales y la priorización cartográfica y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Fotheringham et al., 2002).

El resultado cuantitativo se convierte en una decisión operativa cuando permite programar inspecciones y obras con evidencia territorial. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En tramos de Chone–Flavio Alfaro y redes locales de Manabí, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Moreno-Ponce et al., 2025; Oshan et al., 2019).

2.7.3 Comunicación cartográfica de incertidumbre y exposición

La evaluación se sustenta en superficies de riesgo, exposición, infraestructura y costo, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Deressa et al., 2025; Moreno-Ponce et al., 2025).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante los escenarios espaciales y la priorización cartográfica se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. La evaluación se centra en desempeño, estabilidad y posibilidad de

actualización, no solo en el valor obtenido durante el entrenamiento (Rey et al., 2023; Moreno-Ponce et al., 2025).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: programar inspecciones y obras con evidencia territorial. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En tramos de Chone–Flavio Alfaro y redes locales de Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Rey et al., 2023; Fotheringham et al., 2002).

2.7.4 Uso de resultados espaciales en programación de obras

La evaluación se sustenta en superficies de riesgo, exposición, infraestructura y costo, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Deressa et al., 2025; Anselin, 1995).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante los escenarios espaciales y la priorización cartográfica se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La distancia sobre la red puede diferir de la euclidiana y una transformación de coordenadas incorrecta puede desplazar eventos hasta otro segmento; ambas situaciones deben comprobarse antes del análisis espacial. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Rey et al., 2023; Oshan et al., 2019).

Ejercicio 2.7. Modelos espaciales predictivos para planificación vial

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con atributos geométricos, exposición, ambiente y siniestros georreferenciados. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Entrene un modelo de probabilidad por segmento y reserve una zona completa para prueba espacial.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: mapa de probabilidad acompañado de precisión, recall y falsos negativos. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Practica Python 2

El siguiente laboratorio implementa limpieza y control espacial con geopandas. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 2. Limpieza y análisis espacial con GeoPandas
# Requisitos: numpy, pandas, geopandas, shapely, matplotlib
import numpy as np
```

```
import pandas as pd
import geopandas as gpd
import matplotlib. pyplot as plt
from shapely. geometry import Point, LineString

rng = np. random. default_rng(12)
# Coordenadas UTM didácticas próximas a Jipijapa; no representan siniestros reales.
road = LineString([(545000, 9850000), (548500, 9852600)])
base = np. linspace(0.05, 0.95, 160)
xy = []
for t in base:
 x = 545000 + 3500*t + rng. normal(0, 55)
 y = 9850000 + 2600*t + rng. normal(0, 55)
 xy. append((x, y))
# Agregar duplicados deliberados para demostrar control de calidad.
xy. extend(xy[:5])

df = pd. DataFrame(xy, columns=["este", "norte"])
df["id_evento"] = np. arange(len(df)) + 1
gdf = gpd. GeoDataFrame(
 df,
 geometry=gpd. points_from_xy(df. este, df. norte),
 crs="EPSG:32717" # WGS84 / UTM zona 17S
)

# 1) remover duplicados exactos por coordenada
gdf = gdf. drop_duplicates(subset=["este", "norte"]). copy()
# 2) distancia al eje vial
axis_gdf = gpd. GeoDataFrame({"id_via":[1]}, geometry=[road], crs=gdf. crs)
gdf["dist_eje_m"] = gdf. geometry. distance(road)
# 3) marcar registros que requieren revisión manual
gdf["revisar"] = gdf["dist_eje_m"] > 120
print(gdf[["id_evento", "dist_eje_m", "revisar"]]. head())

ax = axis_gdf. plot(linewidth=3, figsize=(7, 5))
gdf. plot(ax=ax, column="revisar", categorical=True, markersize=18, legend=True)
ax. set_title("Control de posición de eventos respecto del eje vial")
plt. tight_layout()
plt. show()
```

Preguntas para el laboratorio

- ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?
- ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?
- ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?
- ¿Qué información de campo agregaría para reducir incertidumbre?
- ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

2.8 Referencias del capítulo 2

Anselin, L. (1995). Local indicators of spatial association: LISA. Geographical Analysis, 27(2), 93–115. https://doi. org/10.1111/j.1538-4632.1995. tb00338. x

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). Geographically weighted regression: The analysis of spatially varying relationships. Wiley.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. Geographical Analysis, 24(3), 189–206.

Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. ISPRS International Journal of Geo-Information, 8(6), 269. https://doi. org/10.3390/ijgi8060269

Rey, S. J., Arribas-Bel, D., & Wolf, L. J. (2023). Geographic data science with Python. CRC Press.

Deressa, B. F., Habtegiogis, K. A., Endashaw, D. K., Al-Ramadan, B. M., & Al-Ahmadi, H. M. (2025). A systematic review of GIS-driven road traffic accident evaluation. Vehicles, 7(4), 161. https://doi. org/10.3390/vehicles7040161

Skaug, L., Nojoumian, M., Dang, N., & Yap, A. (2025). Road crash analysis and modeling: A systematic review of methods, data, and emerging technologies. Applied Sciences, 15(13), 7115. https://doi. org/10.3390/app15137115

Moreno-Ponce, L. A., Pérez-Zuriaga, A. M., & García, A. (2025). Predictive models and GIS for road safety: Application to a segment of the Chone–Flavio Alfaro road. Sustainability, 17(11), 5032. https://doi. org/10.3390/su17115032

Ponce Quiroz, T. A. (2021). Determinación del nivel de servicio en el tramo de la carretera de dos carriles Jipijapa – Joa [Trabajo de titulación, Universidad Estatal del Sur de Manabí]. Repositorio Digital UNESUM. http://repositorio. unesum. edu. ec/handle/53000/2967

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO III

Modelos de Densidad Kernel y Predicción de Zonas Críticas

AUTOR: Carvajal Rivadeneira Daniel David

EDITORIAL RUNAIKI

3 Modelos de Densidad Kernel y Predicción de Zonas Críticas

Las concentraciones de siniestros no deben definirse por apreciación visual. La densidad Kernel, los estadísticos locales y los modelos de conteo permiten diferenciar acumulaciones relevantes de variaciones aleatorias. Su aplicación exige controlar el ancho de banda, la exposición y la segmentación de la vía. La identificación de concentraciones debe distinguir densidad descriptiva, significancia local y predicción de frecuencia. Combinar KDE con estadística espacial y modelos de conteo reduce la dependencia de un único mapa de calor (Bíl et al., 2013; Xie & Yan, 2013; Moreno-Ponce et al., 2025).

Resultados de aprendizaje

- Explicar la estimación de densidad Kernel y el papel del ancho de banda en la identificación de zonas de concentración.
- Diferenciar concentración visual, significación espacial y frecuencia esperada de siniestros.
- Ajustar modelos Poisson y Binomial Negativa considerando exposición y sobredispersión.
- Interpretar resultados bayesianos espaciales sin confundir probabilidad con causalidad.
- Diseñar reglas reproducibles para seguimiento de zonas críticas emergentes.

3.1 Fundamentos matemáticos del Kernel Density

El análisis de fundamentos matemáticos del kernel density parte de una definición explícita de la unidad de análisis y de la pregunta técnica que se pretende responder. Se consideran coordenadas de eventos, distancia y ancho de banda, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Cameron & Trivedi, 2013; Moreno-Ponce et al., 2025).

El procedimiento adoptado emplea estimación Kernel continua y sobre red y vincula cada salida con una acción verificable: representar intensidad sin confundir concentración con causalidad. La aplicación en siniestros

distribuidos a lo largo de un corredor obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (Hilbe, 2011; Xie & Yan, 2013).

3.1.1 Función Kernel y estimación de intensidad espacial

La evaluación se sustenta en coordenadas de eventos, distancia y ancho de banda, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anderson, 2009; Xie & Yan, 2013).

$$\hat{f}_h(s) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{\| s - s_i \|}{h}\right)$$

La densidad estimada depende del ancho de banda h, del número de eventos n y de la función Kernel K. El valor de h controla el equilibrio entre detalle local y suavizamiento de la superficie.

La estimación se obtiene mediante la estimación Kernel continua y sobre red y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Cameron & Trivedi, 2013).

La utilidad del análisis se hace evidente cuando permite representar intensidad sin confundir concentración con causalidad. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En siniestros distribuidos a lo largo de un corredor, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el

estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2025; Anderson, 2009).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Bíl et al., 2013; Xie & Yan, 2013).

3.1.2 Ancho de banda como parámetro de suavizamiento

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean coordenadas de eventos, distancia y ancho de banda y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Cameron & Trivedi, 2013; Bíl et al., 2013).

La interpretación requiere una comprobación externa; la salida obtenida mediante la estimación Kernel continua y sobre red se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Xie & Yan, 2013; Cameron & Trivedi, 2013).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de representar intensidad sin confundir concentración con causalidad. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En siniestros distribuidos a lo largo de un corredor, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la

medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Xie & Yan, 2013; Hilbe, 2011).

La sensibilidad del resultado se estudia mediante variaciones controladas de los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Hilbe, 2011; Anderson, 2009).

3.1.3 Corrección de borde en corredores y límites administrativos

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de coordenadas de eventos, distancia y ancho de banda se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Moreno-Ponce et al., 2025; Anderson, 2009).

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante la estimación Kernel continua y sobre red frente a una referencia ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Hilbe, 2011; Bíl et al., 2013).

3.1.4 Densidad Kernel sobre redes viales frente a superficies planas

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran coordenadas de eventos, distancia y ancho de banda, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente,

se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Anderson, 2009; Cameron & Trivedi, 2013).

La estimación utiliza estimación Kernel continua y sobre red, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Moreno-Ponce et al., 2025; Anderson, 2009).

La interpretación técnica orienta una acción verificable: representar intensidad sin confundir concentración con causalidad. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En siniestros distribuidos a lo largo de un corredor, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Moreno-Ponce et al., 2025; Bíl et al., 2013).

Ejercicio 3.1. Fundamentos matemáticos de Kernel Density

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con coordenadas de siniestros, severidad y longitud de red. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule manualmente el aporte de un núcleo gaussiano a una celda y luego reproduzca el cálculo para todos los puntos.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: superficie de densidad expresada con unidad y ancho de banda documentados. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

3.2 Parámetros de suavizamiento y calibración del modelo

El análisis de parámetros de suavizamiento y calibración del modelo parte de una definición explícita de la unidad de análisis y de la pregunta técnica que se pretende responder. Se consideran ancho de banda, densidad local y estabilidad de máximos, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Moreno-Ponce et al., 2025; Xie & Yan, 2013).

El procedimiento adoptado emplea calibración, validación y sensibilidad y vincula cada salida con una acción verificable: evitar hotspots que solo aparecen por una elección arbitraria del suavizamiento. La aplicación en redes con densidades de observación desiguales obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (Anderson, 2009; Moreno-Ponce et al., 2025).

3.2.1 Selección basada en criterio estadístico del ancho de banda

La evaluación se sustenta en ancho de banda, densidad local y estabilidad de máximos, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias,

resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hilbe, 2011; Cameron & Trivedi, 2013).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la calibración, la validación y el análisis de sensibilidad se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Cameron & Trivedi, 2013; Moreno-Ponce et al., 2025).

El resultado cuantitativo se convierte en una decisión operativa cuando permite evitar hotspots que solo aparecen por una elección arbitraria del suavizamiento. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En redes con densidades de observación desiguales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Cameron & Trivedi, 2013; Anderson, 2009).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Moreno-Ponce et al., 2025; Xie & Yan, 2013).

3.2.2 Estabilidad de zonas críticas ante cambios de escala

El desarrollo del apartado requiere trabajar con ancho de banda, densidad local y estabilidad de máximos. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anderson, 2009; Xie & Yan, 2013).

El procedimiento combina la calibración, la validación y el análisis de sensibilidad con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Hilbe, 2011).

La interpretación se orienta a una decisión técnica definida: evitar hotspots que solo aparecen por una elección arbitraria del suavizamiento. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En redes con densidades de observación desiguales, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Moreno-Ponce et al., 2025; Anderson, 2009).

Para visualizar sensibilidad de la intensidad kernel ante diferentes anchos de banda, la Figura 3.1 presenta una representación técnica que complementa la explicación y facilita su reproducción.

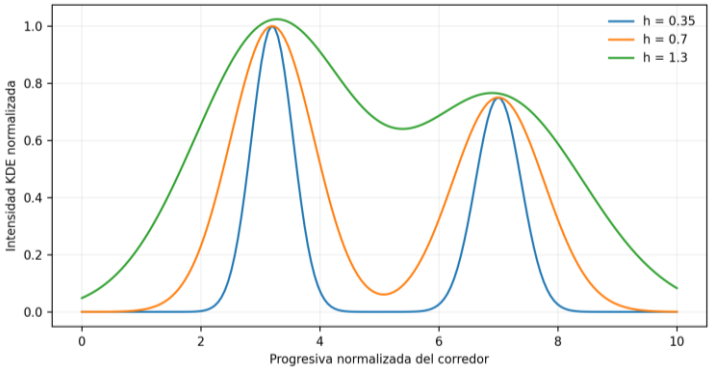


Figura 3.1. Sensibilidad de la intensidad Kernel ante diferentes anchos de banda.

3.2.3 Anchos adaptativos para densidades de datos desiguales

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de ancho de banda, densidad local y estabilidad de máximos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Bíl et al., 2013; Xie & Yan, 2013).

El procedimiento combina la calibración, la validación y el análisis de sensibilidad con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Anderson, 2009; Cameron & Trivedi, 2013).

3.2.4 Pruebas de sensibilidad antes de declarar un hotspot

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de ancho de banda, densidad local y estabilidad de máximos se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Cameron & Trivedi, 2013; Bíl et al., 2013).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora la calibración, la validación y el análisis de sensibilidad. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. El criterio de selección se define antes de

observar el conjunto reservado para prueba (Xie & Yan, 2013; Cameron & Trivedi, 2013).

El resultado cuantitativo se convierte en una decisión operativa cuando permite evitar hotspots que solo aparecen por una elección arbitraria del suavizamiento. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En redes con densidades de observación desiguales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Xie & Yan, 2013; Hilbe, 2011).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Hilbe, 2011; Anderson, 2009).

Ejercicio 3.2. Calibración del ancho de banda

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con coordenadas de siniestros y varios valores de ancho de banda. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Evalúe anchos de banda de 100, 250, 500 y 1000 m; compare estabilidad espacial y capacidad para separar concentraciones cercanas.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.

5) Traduzca el resultado. El producto mínimo será: selección razonada del ancho de banda y gráfico de sensibilidad. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

3.3 Generación computacional de mapas de calor de accidentes

El estudio de generación computacional de mapas de calor de accidentes exige ordenar primero la evidencia y seleccionar después el algoritmo. La base reúne resolución, extensión, exposición y escala de visualización y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Xie & Yan, 2013; Cameron & Trivedi, 2013).

Desde una perspectiva operativa, rasterización y normalización ayuda a estructurar la información y a producir mapas comparables y reproducibles. El análisis se plantea para cartografía de seguridad vial para técnicos y autoridades, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Cameron & Trivedi, 2013; Moreno-Ponce et al., 2025).

3.3.1 Definición de malla, resolución y extensión cartográfica

La evaluación se sustenta en resolución, extensión, exposición y escala de visualización, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anderson, 2009; Xie & Yan, 2013).

La estimación se obtiene mediante rasterización y normalización y se acompaña de una comprobación ingenieril independiente. El contraste entre

resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Cameron & Trivedi, 2013).

La utilidad del análisis se hace evidente cuando permite producir mapas comparables y reproducibles. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En cartografía de seguridad vial para técnicos y autoridades, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2025; Anderson, 2009).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Bíl et al., 2013; Xie & Yan, 2013).

3.3.2 Normalización por exposición para evitar mapas engañosos

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean resolución, extensión, exposición y escala de visualización y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2025; Bíl et al., 2013).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando rasterización y normalización como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta

el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. El criterio de selección se define antes de observar el conjunto reservado para prueba (Hilbe, 2011; Xie & Yan, 2013).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de producir mapas comparables y reproducibles. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En cartografía de seguridad vial para técnicos y autoridades, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Hilbe, 2011; Moreno-Ponce et al., 2025).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Anderson, 2009; Bíl et al., 2013).

3.3.3 Clasificación de intensidades y legibilidad del mapa

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean resolución, extensión, exposición y escala de visualización y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Bíl et al., 2013; Cameron & Trivedi, 2013).

La estimación utiliza rasterización y normalización, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras

que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Anderson, 2009; Hilbe, 2011).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de producir mapas comparables y reproducibles. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En cartografía de seguridad vial para técnicos y autoridades, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Anderson, 2009; Bíl et al., 2013).

Para ordenar parámetros que deben acompañar un mapa kde, la Tabla 3.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 3-1. Parámetros que deben acompañar un mapa KDE

Parámetro	**Efecto**	**Registro mínimo**
Ancho de banda	Suavizamiento	Valor y unidad
Tamaño de celda	Resolución	Metros/píxel
Kernel	Forma de ponderación	Tipo empleado
Extensión	Área analizada	Límites
Exposición	Normalización	Denominador

Nota. Elaboración propia con base en Anderson (2009) y Bíl et al. (2013).

3.3.4 Reproducibilidad de mapas de calor mediante parámetros registrados

El desarrollo del apartado requiere trabajar con resolución, extensión, exposición y escala de visualización. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de

corrección aplicada a valores imposibles o incompletos (Xie & Yan, 2013; Moreno-Ponce et al., 2025).

La estimación utiliza rasterización y normalización, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Bíl et al., 2013; Xie & Yan, 2013).

Ejercicio 3.3. Generación computacional de mapas de calor

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con coordenadas UTM, severidad y tipo de colisión. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Genere una malla regular, calcule KDE con Python y superponga la red vial; exporte el resultado a GeoTIFF o imagen de alta resolución.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: mapa de calor reproducible con escala y sistema de coordenadas. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.

6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

3.4 Identificación de zonas críticas mediante análisis espacial

La comprensión de identificación de zonas críticas mediante análisis espacial requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran estadísticos locales, pesos espaciales y significación; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Xie & Yan, 2013; Cameron & Trivedi, 2013).

La sección articula Getis-Ord Gi* y LISA con una finalidad profesional concreta: separar concentraciones estadísticamente consistentes de variaciones aleatorias. En priorización de zonas críticas sobre redes viales, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Cameron & Trivedi, 2013; Moreno-Ponce et al., 2025).

3.4.1 Umbrales estadísticos para reconocer zonas críticas

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de estadísticos locales, pesos espaciales y significación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Moreno-Ponce et al., 2025; Anderson, 2009).

La validación exige contrastar la salida obtenida mediante Getis-Ord Gi* y LISA con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la

superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Hilbe, 2011; Anderson, 2009).

En términos de ingeniería, el análisis orienta la decisión de separar concentraciones estadísticamente consistentes de variaciones aleatorias. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En priorización de zonas críticas sobre redes viales, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Hilbe, 2011; Xie & Yan, 2013).

3.4.2 Getis-Ord Gi* y concentración local de valores altos

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean estadísticos locales, pesos espaciales y significación y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Cameron & Trivedi, 2013; Bíl et al., 2013).

La validación exige contrastar la salida obtenida mediante Getis-Ord Gi* y LISA con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Xie & Yan, 2013; Cameron & Trivedi, 2013).

En términos de ingeniería, el análisis orienta la decisión de separar concentraciones estadísticamente consistentes de variaciones aleatorias. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En priorización de zonas críticas sobre redes viales, se recomienda

informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Xie & Yan, 2013; Hilbe, 2011).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Hilbe, 2011; Anderson, 2009).

3.4.3 LISA y diferenciación entre agrupamientos y valores atípicos

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean estadísticos locales, pesos espaciales y significación y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Cameron & Trivedi, 2013; Bíl et al., 2013).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante Getis-Ord Gi* y LISA se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Xie & Yan, 2013; Cameron & Trivedi, 2013).

La aplicación del procedimiento cobra sentido al apoyar la decisión de separar concentraciones estadísticamente consistentes de variaciones aleatorias. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En priorización de zonas críticas sobre redes viales, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Xie & Yan, 2013; Hilbe, 2011).

3.4.4 Persistencia espacial y control de falsos positivos

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de estadísticos locales, pesos espaciales y significación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Cameron & Trivedi, 2013; Bíl et al., 2013).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora Getis-Ord Gi* y LISA. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. El criterio de selección se define antes de observar el conjunto reservado para prueba (Xie & Yan, 2013; Cameron & Trivedi, 2013).

El resultado cuantitativo se convierte en una decisión operativa cuando permite separar concentraciones estadísticamente consistentes de variaciones aleatorias. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En priorización de zonas críticas sobre redes viales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Xie & Yan, 2013; Hilbe, 2011).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Hilbe, 2011; Anderson, 2009).

Ejercicio 3.4. Identificación de zonas de concentración de siniestros

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con densidad KDE, número de siniestros, severidad y exposición. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Defina umbrales mediante percentiles y verifique que las zonas seleccionadas también presenten exposición o severidad relevante.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: clasificación alta-media-baja que no dependa únicamente del número bruto de eventos. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

3.5 Modelación estadística de frecuencia de accidentes: Poisson y Binomial Negativa

El estudio de modelación estadística de frecuencia de accidentes: poisson y binomial negativa exige ordenar primero la evidencia y seleccionar después el algoritmo. La base reúne conteo de siniestros, exposición y covariables viales y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Hilbe, 2011; Xie & Yan, 2013).

Desde una perspectiva operativa, Poisson, Binomial Negativa y offsets ayuda a estructurar la información y a estimar frecuencia esperada respetando la naturaleza discreta de la respuesta. El análisis se plantea para segmentos con

distintos volúmenes de tránsito, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Moreno-Ponce et al., 2025; Anderson, 2009).

3.5.1 Supuesto de equidispersión en la regresión de Poisson

La evaluación se sustenta en conteo de siniestros, exposición y covariables viales, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anderson, 2009; Xie & Yan, 2013).

$$P(Y_i = y) = \frac{e^{-\lambda_i}\lambda_i^y}{y!}$$

En el modelo de Poisson, la media y la varianza condicional se consideran iguales a λ_i. Esta propiedad debe comprobarse, porque los conteos de siniestros suelen presentar dispersión superior a la prevista por el modelo.

La estimación se obtiene mediante Poisson, Binomial Negativa y offsets y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Cameron & Trivedi, 2013).

La utilidad del análisis se hace evidente cuando permite estimar frecuencia esperada respetando la naturaleza discreta de la respuesta. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En segmentos con distintos volúmenes de tránsito, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de

construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2025; Anderson, 2009).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Bíl et al., 2013; Xie & Yan, 2013).

3.5.2 Sobredispersión y formulación Binomial Negativa

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de conteo de siniestros, exposición y covariables viales se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Cameron & Trivedi, 2013; Hilbe, 2011).

$$\mathrm{Var}(Y_i) = \mu_i + \alpha\mu_i^2$$

La Binomial Negativa incorpora el parámetro de dispersión α. Cuando α es mayor que cero, la varianza puede superar a la media y representar mejor la heterogeneidad presente entre segmentos viales.

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando Poisson, Binomial Negativa y offsets como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. El criterio de selección se define antes de observar el conjunto reservado para prueba (Xie & Yan, 2013; Hilbe, 2011).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de estimar frecuencia esperada respetando la naturaleza discreta de la respuesta. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En

segmentos con distintos volúmenes de tránsito, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Xie & Yan, 2013; Anderson, 2009).

Para visualizar relación media-varianza en modelos poisson y binomial negativa, la Figura 3.2 presenta una representación técnica que complementa la explicación y facilita su reproducción.

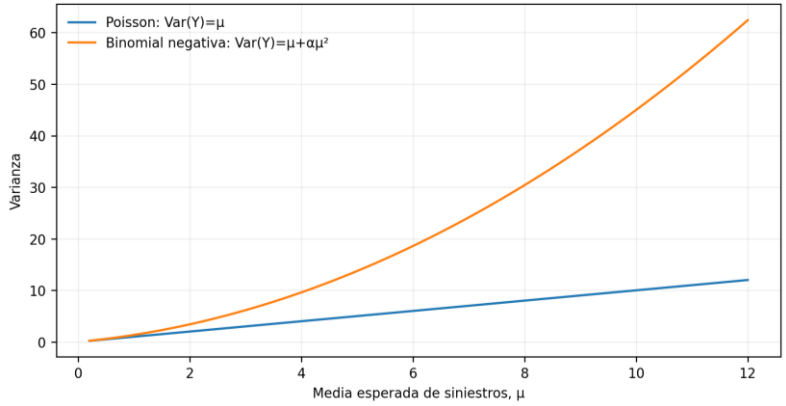


Figura 3.2. Relación media-varianza en modelos Poisson y Binomial Negativa.

Nota. Elaboración propia con base en la formulación de Cameron y Trivedi (2013) e Hilbe (2011).

3.5.3 Offsets de exposición y comparación entre segmentos

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran conteo de siniestros, exposición y covariables viales, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Xie & Yan, 2013; Bíl et al., 2013).

El procedimiento combina Poisson, Binomial Negativa y offsets con una verificación técnica que permita detectar resultados poco plausibles. La

comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Bíl et al., 2013; Cameron & Trivedi, 2013).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de estimar frecuencia esperada respetando la naturaleza discreta de la respuesta. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En segmentos con distintos volúmenes de tránsito, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Bíl et al., 2013; Hilbe, 2011).

Para ordenar diagnóstico para escoger entre poisson y binomial negativa, la Tabla 3.2 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 3-2. Diagnóstico para escoger entre Poisson y Binomial Negativa

Comprobación	**Poisson**	**Binomial Negativa**
Media-varianza	Aproximadamente iguales	Varianza mayor
Dispersión	Cercana a 1	Superior a 1
Parámetro α	No aplica	Estimable
Offset	Admitido	Admitido
Lectura	Conteos regulares	Heterogeneidad extra

Nota. Elaboración propia con base en Cameron y Trivedi (2013) e Hilbe (2011).

3.5.4 Diagnóstico de residuos y capacidad predictiva en conteos

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran conteo de siniestros, exposición y covariables viales, diferenciando los

atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Hilbe, 2011; Xie & Yan, 2013).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora Poisson, Binomial Negativa y offsets. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. El criterio de selección se define antes de observar el conjunto reservado para prueba (Cameron & Trivedi, 2013; Hilbe, 2011).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: estimar frecuencia esperada respetando la naturaleza discreta de la respuesta. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En segmentos con distintos volúmenes de tránsito, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Cameron & Trivedi, 2013; Moreno-Ponce et al., 2025).

Ejercicio 3.5. Modelación de frecuencia con Poisson y Binomial Negativa

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con conteo de siniestros, TPDA, longitud, V85 y accesos. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.

3) Ejecute el método. Ajuste Poisson, examine sobredispersión y, si la varianza excede claramente a la media, ajuste Binomial Negativa con offset de exposición.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: comparación AIC, residuos y predicción por segmento. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Ejemplo breve. Ejemplo de decisión: si la media del conteo es 2,1 y la varianza es 8,7, Poisson puede resultar restrictivo; compare Binomial Negativa y examine AIC y residuos.

3.6 Integración de modelos Kernel con enfoques bayesianos espaciales

Para trabajar con integración de modelos kernel con enfoques bayesianos espaciales de forma reproducible, se establece primero la unidad de análisis y la pregunta que orienta el procedimiento. Se consideran tasas observadas, información previa y efectos espaciales, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Bíl et al., 2013; Moreno-Ponce et al., 2025).

El procedimiento adoptado emplea modelos jerárquicos y suavizamiento bayesiano y vincula cada salida con una acción verificable: reducir inestabilidad en áreas con pocos eventos. La aplicación en corredores rurales donde los conteos son escasos obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (Xie & Yan, 2013; Cameron & Trivedi, 2013).

3.6.1 Estructura jerárquica para combinar información previa y observada

En este apartado, la evidencia se organiza a partir de tasas observadas, información previa y efectos espaciales. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anderson, 2009; Bíl et al., 2013).

La evaluación utiliza modelos jerárquicos y suavizamiento bayesiano, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Xie & Yan, 2013).

En términos de ingeniería, el análisis orienta la decisión de reducir inestabilidad en áreas con pocos eventos. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En corredores rurales donde los conteos son escasos, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Moreno-Ponce et al., 2025; Hilbe, 2011).

La estabilidad del resultado se evalúa mediante cambios controlados en los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Bíl et al., 2013; Anderson, 2009).

3.6.2 Efectos aleatorios espaciales y suavizamiento de tasas inestables

La evaluación se sustenta en tasas observadas, información previa y efectos espaciales, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Anderson, 2009; Xie & Yan, 2013).

La estimación se obtiene mediante modelos jerárquicos y suavizamiento bayesiano y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2025; Cameron & Trivedi, 2013).

La utilidad del análisis se hace evidente cuando permite reducir inestabilidad en áreas con pocos eventos. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En corredores rurales donde los conteos son escasos, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2025; Anderson, 2009).

3.6.3 Distribuciones posteriores e intervalos creíbles

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran tasas observadas, información previa y efectos espaciales, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta

forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Anderson, 2009; Bíl et al., 2013).

La estimación utiliza modelos jerárquicos y suavizamiento bayesiano, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Moreno-Ponce et al., 2025; Bíl et al., 2013).

La utilidad del análisis se hace evidente cuando permite reducir inestabilidad en áreas con pocos eventos. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En corredores rurales donde los conteos son escasos, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2025; Cameron & Trivedi, 2013).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Bíl et al., 2013; Moreno-Ponce et al., 2025).

3.6.4 Lectura de mapas probabilísticos para decisiones preventivas

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran tasas observadas, información previa y efectos espaciales, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente,

se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Anderson, 2009; Xie & Yan, 2013).

La estimación utiliza modelos jerárquicos y suavizamiento bayesiano, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Moreno-Ponce et al., 2025; Hilbe, 2011).

Los resultados aportan valor técnico cuando permiten reducir inestabilidad en áreas con pocos eventos. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En corredores rurales donde los conteos son escasos, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2025; Anderson, 2009).

Ejercicio 3.6. Integración de Kernel con estimación bayesiana espacial

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con conteos por segmento, exposición y estimación media de la red. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.

3) Ejecute el método. Calcule una tasa observada y una tasa suavizada por Empirical Bayes; otorgue mayor peso a la media global cuando la exposición sea baja.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: ranking de segmentos antes y después del suavizamiento. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
1) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

3.7 Predicción dinámica de hotspots en redes viales

Antes de modelar predicción dinámica de hotspots en redes viales, es necesario organizar la evidencia y comprobar que responda a la pregunta técnica planteada. La base reúne ventanas temporales, intensidad y persistencia espacial y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Hilbe, 2011; Anderson, 2009).

En términos operativos, la actualización dinámica y las alertas facilita la organización de la evidencia necesaria para detectar sitios emergentes sin reaccionar a fluctuaciones aisladas. El análisis se plantea para seguimiento periódico de corredores de Manabí, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Moreno-Ponce et al., 2025; Xie & Yan, 2013).

3.7.1 Ventanas móviles para seguimiento de concentraciones emergentes

El desarrollo del apartado requiere trabajar con ventanas temporales, intensidad y persistencia espacial. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse,

actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hilbe, 2011; Xie & Yan, 2013).

La interpretación requiere una comprobación externa; la salida obtenida mediante la actualización dinámica y las alertas se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Cameron & Trivedi, 2013; Hilbe, 2011).

Los resultados aportan valor técnico cuando permiten detectar sitios emergentes sin reaccionar a fluctuaciones aisladas. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En seguimiento periódico de corredores de Manabí, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Cameron & Trivedi, 2013; Moreno-Ponce et al., 2025).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Moreno-Ponce et al., 2025; Bíl et al., 2013).

3.7.2 Actualización periódica de intensidad y frecuencia esperada

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran ventanas temporales, intensidad y persistencia espacial, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra

la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Hilbe, 2011; Xie & Yan, 2013).

La aplicación metodológica de la actualización dinámica y las alertas sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. El criterio de selección se define antes de observar el conjunto reservado para prueba (Cameron & Trivedi, 2013; Hilbe, 2011).

La predicción dinámica de zonas críticas exige representar el resultado esperado para un periodo posterior y no únicamente el mapa histórico. La Figura 3.3 muestra una estimación espacial de intensidad para el siguiente periodo mediante una superficie Kernel construida con eventos simulados y una actualización temporal sencilla.

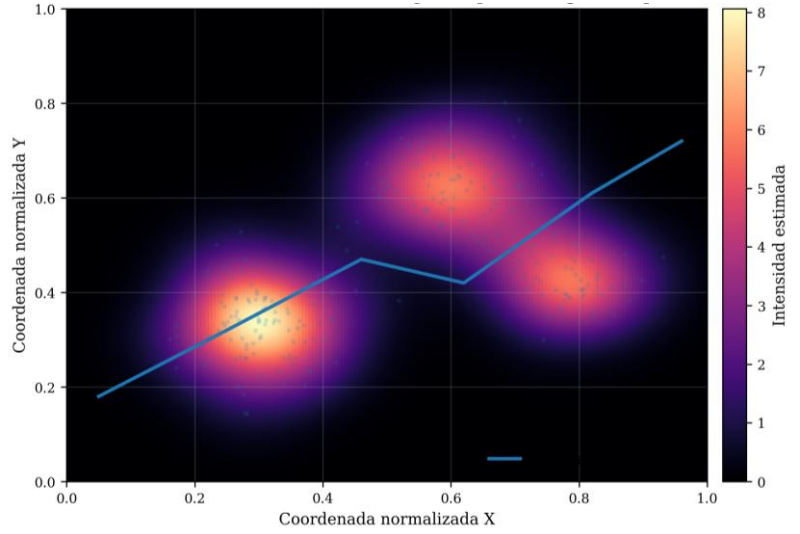


Figura 3.3. Intensidad espacial predicha para el siguiente periodo de análisis.

Nota. Elaboración propia con eventos simulados del Laboratorio Python 3.

La concentración estimada permite identificar sectores donde la inspección puede programarse con mayor prioridad. Una zona de alta intensidad no constituye por sí sola una relación causal ni una orden de obra; debe verificarse su persistencia, exposición y coherencia con las características físicas de la red.

3.7.3 Criterios de alerta temprana para inspección de corredores

Para abordar el problema se integran ventanas temporales, intensidad y persistencia espacial, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hilbe, 2011; Bíl et al., 2013).

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante la actualización dinámica y las alertas frente a una referencia ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Cameron & Trivedi, 2013; Hilbe, 2011).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de detectar sitios emergentes sin reaccionar a fluctuaciones aisladas. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En seguimiento periódico de corredores de Manabí, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Cameron & Trivedi, 2013; Moreno-Ponce et al., 2025).

3.7.4 Evaluación retrospectiva de alertas y estabilidad espacial

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de ventanas temporales, intensidad y persistencia espacial se examinan mediante rangos

físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Bíl et al., 2013; Cameron & Trivedi, 2013).

Para sostener la interpretación, el procedimiento basado en la actualización dinámica y las alertas se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. Un ancho de banda pequeño puede fragmentar la superficie en máximos inestables, mientras que uno excesivo puede fusionar zonas distintas; la comparación debe repetirse con exposición y borde controlados. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Anderson, 2009; Hilbe, 2011).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: detectar sitios emergentes sin reaccionar a fluctuaciones aisladas. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En seguimiento periódico de corredores de Manabí, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Anderson, 2009; Bíl et al., 2013).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Xie & Yan, 2013; Cameron & Trivedi, 2013).

Ejercicio 3.7. Predicción dinámica de zonas de alta concentración

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con fecha, coordenadas, severidad y segmento. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Aplique ventanas móviles de 12 meses y recalcule KDE o conteos por segmento; mida persistencia de cada zona prioritaria.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: tabla de aparición, permanencia y desaparición de zonas de alta concentración. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Practica Python 3

El siguiente laboratorio implementa kde y binomial negativa. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 3. KDE espacial y conteos con Binomial Negativa
# Requisitos: numpy, pandas, scikit-learn, statsmodels, matplotlib
import numpy as np
import pandas as pd
import matplotlib. pyplot as plt
from sklearn. neighbors import KernelDensity
import statsmodels. api as sm
import statsmodels. formula. api as smf

rng = np. random. default_rng(18)
```

```
centros = np. array([[0.25, 0.30], [0.55, 0.62], [0.78, 0.38]])
puntos = np. vstack([
 c + rng. normal(0, sd, size=(n, 2))
 for c, n, sd in zip(centros, [90, 65, 45], [0.07, 0.06, 0.05])
])
puntos = np. clip(puntos, 0, 1)

# KDE
malla = np. linspace(0, 1, 160)
xx, yy = np. meshgrid(malla, malla)
grid = np. c_[xx. ravel(), yy. ravel()]
kde = KernelDensity(kernel="gaussian", bandwidth=0.065). fit(puntos)
densidad = np. exp(kde. score_samples(grid)). reshape(xx. shape)

plt. figure(figsize=(7, 5))
plt. imshow(densidad, origin="lower", extent=[0,1,0,1], aspect="auto")
plt. scatter(puntos[:,0], puntos[:,1], s=4, alpha=0.25)
plt. colorbar(label="Densidad relativa")
plt. title("KDE de eventos")
plt. tight_layout()
plt. show()

# Conteos por segmento y Binomial Negativa
n = 180
segmentos = pd. DataFrame({
 "long_km": rng. uniform(0.2, 1.5, n),
 "tpda": rng. lognormal(np. log(650), 0.6, n),
 "curvatura": rng. exponential(0.004, n),
 "berma": rng. integers(0, 2, n)
})
mu = np. exp(-2.1 + 0.00035*segmentos. tpda + 38*segmentos. curvatura
 - 0.35*segmentos. berma + np. log(segmentos. long_km))
# Gamma-Poisson para introducir sobredispersión.
alpha = 0.8
lam = rng. gamma(shape=1/alpha, scale=mu*alpha)
segmentos["siniestros"] = rng. poisson(lam)

modelo = smf. glm(
 "siniestros ~ tpda + curvatura + berma",
 data=segmentos,
 family=sm. families. NegativeBinomial(alpha=alpha),
 offset=np. log(segmentos["long_km"])
). fit()
print(modelo. summary())
```

Preguntas para el laboratorio

• ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?

• ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?

• ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?

• ¿Qué información de campo agregaría para reducir incertidumbre?

• ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

3.8 Referencias del capítulo 3

Anderson, T. K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. Accident Analysis & Prevention, 41(3), 359–364. https://doi. org/10.1016/j. aap.2008.12.014

Bíl, M., Andrášik, R., & Janoška, Z. (2013). Identification of hazardous road locations of traffic accidents by means of kernel density estimation and cluster significance evaluation. Accident Analysis & Prevention, 55, 265–273. https://doi. org/10.1016/j. aap.2013.03.003

Xie, Z., & Yan, J. (2013). Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: An integrated approach. Journal of Transport Geography, 31, 64–71. https://doi. org/10.1016/j. jtrangeo.2013.05.009

Cameron, A. C., & Trivedi, P. K. (2013). Regression analysis of count data (2nd ed.). Cambridge University Press.

Hilbe, J. M. (2011). Negative binomial regression (2nd ed.). Cambridge University Press.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. Geographical Analysis, 24(3), 189–206.

Moreno-Ponce, L. A., Pérez-Zuriaga, A. M., & García, A. (2025). Predictive models and GIS for road safety: Application to a segment of the Chone–Flavio Alfaro road. Sustainability, 17(11), 5032. https://doi. org/10.3390/su17115032

Deressa, B. F., Habtegiogis, K. A., Endashaw, D. K., Al-Ramadan, B. M., & Al-Ahmadi, H. M. (2025). A systematic review of GIS-driven road traffic accident evaluation. Vehicles, 7(4), 161. https://doi. org/10.3390/vehicles7040161

Skaug, L., Nojoumian, M., Dang, N., & Yap, A. (2025). Road crash analysis and modeling: A systematic review of methods, data, and emerging technologies. Applied Sciences, 15(13), 7115. https://doi. org/10.3390/app15137115

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO IV

Modelos Temporales ARIMA en la Predicción de Siniestralidad

AUTOR: Delgado Alvia Ronald Pastor

EDITORIAL RUNAIKI

4 Modelos Temporales ARIMA en la Predicción de Siniestralidad

Las series de siniestros contienen dependencia temporal, ciclos estacionales, cambios de nivel y episodios extraordinarios. ARIMA y sus extensiones ofrecen una base transparente para pronósticos de corto plazo, especialmente cuando la cantidad de observaciones no justifica redes profundas. La comparación entre SARIMA, Prophet y LSTM muestra la conveniencia de utilizar una misma partición temporal y métricas homogéneas. Esta lógica también ha sido aplicada en un corredor de Manabí con registros oficiales de siniestros (Feng et al., 2022; Moreno-Ponce et al., 2025).

Resultados de aprendizaje

- Preparar series temporales de seguridad vial conservando el orden real de la información.
- Identificar nivel, estacionalidad y estacionariedad antes del ajuste de modelos ARIMA.
- Seleccionar órdenes p, d y q mediante ACF, PACF y criterios de información.
- Validar pronósticos mediante origen rodante, diagnóstico residual y métricas por horizonte.
- Extender el análisis hacia SARIMA y ARIMAX cuando la evidencia temporal lo justifique.

4.1 Series temporales en seguridad vial y estructuración de datos

El estudio de series temporales en seguridad vial y estructuración de datos exige ordenar primero la evidencia y seleccionar después el algoritmo. La base reúne fecha, periodo, frecuencia de agregación y revisiones de registro y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Feng et al., 2022; Moreno-Ponce et al., 2025).

La aplicación a series oficiales de siniestros de Ecuador requiere una serie cronológica coherente, datos versionados y reglas de decisión establecidas antes del ajuste. La prioridad consiste en preservar el orden real de la

información. El desempeño estadístico se valora únicamente cuando mejora esa decisión. El resultado adquiere valor práctico cuando puede contrastarse con evidencia independiente, ya provenga de inspección, normativa, medición de campo o un periodo posterior (Hamilton, 1994; Hyndman & Athanasopoulos, 2021).

4.1.1 Fecha del evento, fecha de registro y orden temporal

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran fecha, periodo, frecuencia de agregación y revisiones de registro, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Box et al., 2015; Moreno-Ponce et al., 2025).

La estimación utiliza construcción de una serie cronológica coherente, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Instituto Nacional de Estadística y Censos, 2025; Box et al., 2015).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: preservar el orden real de la información. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series oficiales de siniestros de Ecuador, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Instituto Nacional de Estadística y Censos, 2025; Hyndman & Athanasopoulos, 2021).

4.1.2 Frecuencia de agregación y consistencia del calendario

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de fecha, periodo, frecuencia de agregación y revisiones de registro se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Instituto Nacional de Estadística y Censos, 2025; Moreno-Ponce et al., 2025).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la construcción de una serie cronológica coherente se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Moreno-Ponce et al., 2025; Box et al., 2015).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: preservar el orden real de la información. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series oficiales de siniestros de Ecuador, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Moreno-Ponce et al., 2025; Hyndman & Athanasopoulos, 2021).

4.1.3 Tratamiento de meses sin datos, faltantes y revisiones administrativas

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran fecha, periodo, frecuencia de agregación y revisiones de registro, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la

transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Moreno-Ponce et al., 2025; Box et al., 2015).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora la construcción de una serie cronológica coherente. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El criterio de selección se define antes de observar el conjunto reservado para prueba (Hamilton, 1994; Feng et al., 2022).

La utilidad del análisis se hace evidente cuando permite preservar el orden real de la información. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En series oficiales de siniestros de Ecuador, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Hamilton, 1994; Moreno-Ponce et al., 2025).

Para ordenar control de calidad antes de modelar una serie temporal, la Tabla 4.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 4-1. Control de calidad antes de modelar una serie temporal

Comprobación	Pregunta de control	Acción
Frecuencia	¿Todos los periodos existen?	Completar calendario
Faltantes	¿Cero significa ausencia o falta?	Distinguir estados
Revisiones	¿Cambió un valor histórico?	Versionar
Rupturas	¿Hubo cambio operativo?	Documentar intervención

Orden	¿Existe información futura?	Separar antes del ajuste

Nota. Elaboración propia a partir de principios de Box et al. (2015) y Hyndman y Athanasopoulos (2021).

4.1.4 Rupturas estructurales asociadas con cambios operativos o normativos

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran fecha, periodo, frecuencia de agregación y revisiones de registro, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Box et al., 2015; Feng et al., 2022).

La estimación utiliza construcción de una serie cronológica coherente, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Instituto Nacional de Estadística y Censos, 2025; Moreno-Ponce et al., 2025).

El resultado es útil cuando permite preservar el orden real de la información. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En series oficiales de siniestros de Ecuador, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Instituto Nacional de Estadística y Censos, 2025; Box et al., 2015).

Ejercicio 4.1. Series temporales en seguridad vial

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con fecha mensual, siniestros, fallecidos, lesionados y exposición. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Construya una serie regular, complete meses ausentes solo cuando la fuente permita distinguir cero de dato faltante y reserve los últimos 12 meses.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: serie lista para modelación con partición temporal explícita. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

4.2 Componentes de trayectoria, estacionalidad y ruido en siniestralidad

La ingeniería civil aborda componentes de trayectoria, estacionalidad y ruido en siniestralidad mediante la integración de evidencia física y análisis cuantitativo, no mediante una rutina estadística aislada. El análisis combina nivel, estacionalidad, variación aleatoria y cambios de régimen y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá

inspeccionarse (Instituto Nacional de Estadística y Censos, 2025; Hyndman & Athanasopoulos, 2021).

En series mensuales o trimestrales de seguridad vial, la descomposición y las pruebas de estacionariedad deben ejecutarse con parámetros documentados y datos versionados. La finalidad técnica es decidir transformaciones antes del ajuste; el valor de una métrica se interpreta en función de esa finalidad. La decisión debe apoyarse en una verificación independiente del ajuste, mediante inspección, referencia normativa, medición adicional o datos posteriores (Box et al., 2015; Hamilton, 1994).

4.2.1 Nivel de la serie y crecimiento de largo plazo

Para abordar el problema se integran nivel, estacionalidad, variación aleatoria y cambios de régimen, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Box et al., 2015; Hamilton, 1994).

Para sostener la interpretación, el procedimiento basado en descomposición y pruebas de estacionariedad se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Instituto Nacional de Estadística y Censos, 2025; Box et al., 2015).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de decidir transformaciones antes del ajuste. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En series mensuales o trimestrales de seguridad vial, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Instituto Nacional de Estadística y Censos, 2025; Hyndman & Athanasopoulos, 2021).

Como verificación complementaria, se propone retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Hyndman & Athanasopoulos, 2021; Hamilton, 1994).

4.2.2 Estacionalidad mensual, trimestral y efectos de calendario

En este apartado, la evidencia se organiza a partir de nivel, estacionalidad, variación aleatoria y cambios de régimen. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Feng et al., 2022; Hamilton, 1994).

La estimación utiliza descomposición y pruebas de estacionariedad, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Hyndman & Athanasopoulos, 2021; Moreno-Ponce et al., 2025).

En términos de ingeniería, el análisis orienta la decisión de decidir transformaciones antes del ajuste. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En series mensuales o trimestrales de seguridad vial, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Hyndman & Athanasopoulos, 2021; Box et al., 2015).

Para visualizar componentes ilustrativos de una serie temporal de siniestralidad, la Figura 4.1 presenta una representación técnica que complementa la explicación y facilita su reproducción.

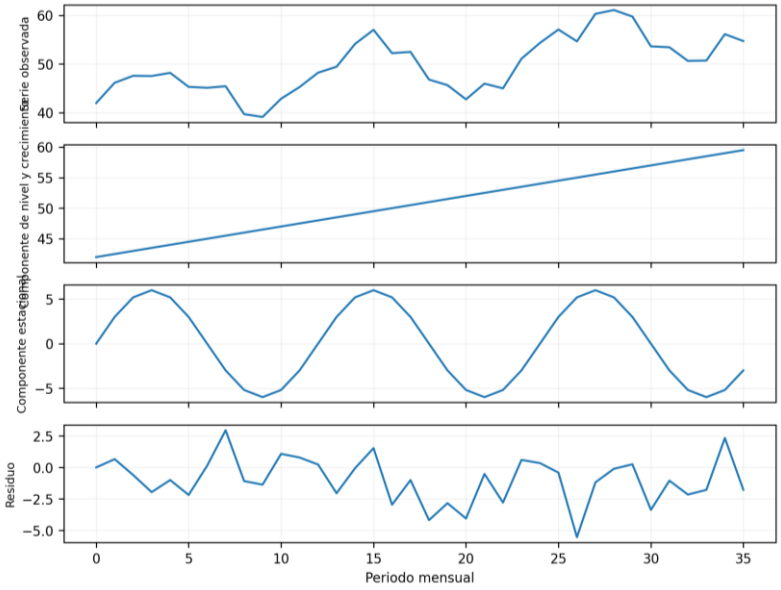


Figura 4.1. Componentes ilustrativos de una serie temporal de siniestralidad.

Nota. Elaboración propia. La serie se genera con fines explicativos para separar nivel, estacionalidad y residuo.

4.2.3 Estacionariedad y transformaciones antes del ajuste

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de nivel, estacionalidad, variación aleatoria y cambios de régimen se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Hyndman & Athanasopoulos, 2021; Hamilton, 1994).

La evaluación utiliza descomposición y pruebas de estacionariedad, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Box et al., 2015; Instituto Nacional de Estadística y Censos, 2025).

4.2.4 Separación entre señal persistente y variación aleatoria

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean nivel, estacionalidad, variación aleatoria y cambios de régimen y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Hyndman & Athanasopoulos, 2021; Instituto Nacional de Estadística y Censos, 2025).

La estimación utiliza descomposición y pruebas de estacionariedad, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de decidir transformaciones antes del ajuste. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En series mensuales o trimestrales de seguridad vial, se recomienda informar el valor central, un rango de incertidumbre, la fuente

principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Box et al., 2015; Feng et al., 2022).

Ejercicio 4.2. Componentes temporales de la siniestralidad

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con serie mensual de siniestros de al menos 48 observaciones. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Descomponga nivel, componente estacional y residuo mediante media móvil o STL; compare amplitud estacional entre años.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: gráfico de componentes y explicación de qué parte puede ser pronosticada. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

4.3 Formulación matemática del modelo ARIMA

Para trabajar con formulación matemática del modelo arima de forma reproducible, se establece primero la unidad de análisis y la pregunta que orienta el procedimiento. Se consideran rezagos, diferencias y errores pasados, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Hamilton, 1994; Instituto Nacional de Estadística y Censos, 2025).

$$\Phi(B)(1-B)^d y_t = c + \Theta(B)\varepsilon_t$$

La formulación reúne el componente autorregresivo Φ(B), la diferenciación de orden d y el componente de medias móviles Θ(B). El operador B desplaza la serie en el tiempo y permite expresar de forma compacta la dependencia temporal.

La sección articula ARIMA y operadores de rezago con una finalidad profesional concreta: representar dependencia temporal con una estructura parsimoniosa. En pronósticos de frecuencia de siniestros, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Moreno-Ponce et al., 2025; Hyndman & Athanasopoulos, 2021).

4.3.1 Operadores autorregresivos y memoria de la serie

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean rezagos, diferencias y errores pasados y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Hyndman & Athanasopoulos, 2021; Hamilton, 1994).

La estimación utiliza ARIMA y operadores de rezago, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Box et al., 2015; Moreno-Ponce et al., 2025).

4.3.2 Diferenciación como herramienta para estabilizar el nivel

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran rezagos, diferencias y errores pasados, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Box et al., 2015; Feng et al., 2022).

La estimación utiliza ARIMA y operadores de rezago, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Instituto Nacional de Estadística y Censos, 2025; Moreno-Ponce et al., 2025).

4.3.3 Componente de medias móviles y propagación de errores

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de rezagos, diferencias y errores pasados se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Instituto Nacional de Estadística y Censos, 2025; Moreno-Ponce et al., 2025).

La interpretación requiere una comprobación externa; la salida obtenida mediante ARIMA y operadores de rezago se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La evaluación se centra en desempeño,

estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Moreno-Ponce et al., 2025; Box et al., 2015).

4.3.4 Interpretación conjunta de AR, I y MA en seguridad vial

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de rezagos, diferencias y errores pasados se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Hyndman & Athanasopoulos, 2021; Feng et al., 2022).

La estimación se obtiene mediante ARIMA y operadores de rezago y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Box et al., 2015; Feng et al., 2022).

La aplicación del procedimiento cobra sentido al apoyar la decisión de representar dependencia temporal con una estructura parsimoniosa. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En pronósticos de frecuencia de siniestros, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Box et al., 2015; Moreno-Ponce et al., 2025).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar

una inspección o una programación preliminar (Feng et al., 2022; Box et al., 2015).

Ejercicio 4.3. Formulación matemática del modelo ARIMA

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con serie estacionaria o diferenciada. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Escriba la ecuación ARIMA seleccionada, calcule una predicción de un paso con coeficientes dados y verifiquela con Python.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: predicción manual y computacional coincidentes dentro del redondeo. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

4.4 Identificación de parámetros p, d y q mediante análisis estadístico

La comprensión de identificación de parámetros p, d y q mediante análisis estadístico requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran ACF, PACF, AIC, BIC y orden de diferenciación; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado

mediante SIG o código y una predicción producida por el modelo (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

Para trabajar con series cortas y medianas de registros administrativos, la identificación y comparación de modelos candidatos debe seguir criterios explícitos y utilizar datos versionados. La prioridad consiste en evitar seleccionar órdenes por inspección aislada. El desempeño estadístico se valora únicamente cuando mejora esa decisión. El resultado adquiere valor práctico cuando puede contrastarse con evidencia independiente, ya provenga de inspección, normativa, medición de campo o un periodo posterior (Hyndman & Athanasopoulos, 2021; Hamilton, 1994).

4.4.1 Lectura de ACF para dependencia serial

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de ACF, PACF, AIC, BIC y orden de diferenciación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Instituto Nacional de Estadística y Censos, 2025; Hyndman & Athanasopoulos, 2021).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la identificación y comparación de modelos candidatos se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Moreno-Ponce et al., 2025; Hamilton, 1994).

La aplicación del procedimiento cobra sentido al apoyar la decisión de evitar seleccionar órdenes por inspección aislada. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En series cortas y medianas de registros administrativos, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría

comprobar si la prioridad asignada tiene respaldo en terreno (Moreno-Ponce et al., 2025; Instituto Nacional de Estadística y Censos, 2025).

4.4.2 PACF y selección preliminar del orden autorregresivo

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean ACF, PACF, AIC, BIC y orden de diferenciación y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Instituto Nacional de Estadística y Censos, 2025; Hyndman & Athanasopoulos, 2021).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora la identificación y comparación de modelos candidatos. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El criterio de selección se define antes de observar el conjunto reservado para prueba (Moreno-Ponce et al., 2025; Hamilton, 1994).

El resultado cuantitativo se convierte en una decisión operativa cuando permite evitar seleccionar órdenes por inspección aislada. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En series cortas y medianas de registros administrativos, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Moreno-Ponce et al., 2025; Instituto Nacional de Estadística y Censos, 2025).

Como prueba de robustez, se modifican de forma controlada los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar

una inspección o una programación preliminar (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

4.4.3 AIC y BIC para comparar modelos candidatos

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean ACF, PACF, AIC, BIC y orden de diferenciación y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Hamilton, 1994; Instituto Nacional de Estadística y Censos, 2025).

La interpretación requiere una comprobación externa; la salida obtenida mediante la identificación y comparación de modelos candidatos se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Feng et al., 2022; Hyndman & Athanasopoulos, 2021).

La interpretación técnica orienta una acción verificable: evitar seleccionar órdenes por inspección aislada. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En series cortas y medianas de registros administrativos, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Feng et al., 2022; Hamilton, 1994).

4.4.4 Parsimonia frente a complejidad innecesaria

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran ACF, PACF, AIC, BIC y orden de diferenciación, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin

depender de decisiones implícitas del analista (Box et al., 2015; Instituto Nacional de Estadística y Censos, 2025).

La estimación utiliza identificación y comparación de modelos candidatos, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Instituto Nacional de Estadística y Censos, 2025; Hyndman & Athanasopoulos, 2021).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de evitar seleccionar órdenes por inspección aislada. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En series cortas y medianas de registros administrativos, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Instituto Nacional de Estadística y Censos, 2025; Feng et al., 2022).

Ejercicio 4.4. Identificación de parámetros p, d y q

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con serie mensual y diferencias de primer orden. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule ACF y PACF, proponga tres combinaciones p-d-q y compare AIC sin usar el conjunto de prueba para escoger el modelo.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: tabla de candidatos con justificación estadística. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

4.5 Validación, diagnóstico y ajuste del modelo

En ingeniería civil, validación, diagnóstico y ajuste del modelo debe abordarse como un problema técnico y no como una simple ejecución estadística. El análisis combina residuos, horizontes, errores y cobertura y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Hyndman & Athanasopoulos, 2021; Box et al., 2015).

La sección articula validación de origen rodante y diagnóstico residual con una finalidad profesional concreta: comprobar estabilidad del pronóstico a medida que avanza el tiempo. En evaluación trimestral de modelos de seguridad, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Feng et al., 2022; Hamilton, 1994).

4.5.1 Validación de origen rodante y horizonte de predicción

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de residuos, horizontes, errores y cobertura se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una

condición vial real que precisamente interesa estudiar (Hyndman & Athanasopoulos, 2021; Hamilton, 1994).

El procedimiento combina validación de origen rodante y diagnóstico residual con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Box et al., 2015; Moreno-Ponce et al., 2025).

La interpretación técnica orienta una acción verificable: comprobar estabilidad del pronóstico a medida que avanza el tiempo. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En evaluación trimestral de modelos de seguridad, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

4.5.2 Diagnóstico de residuos y prueba de autocorrelación remanente

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean residuos, horizontes, errores y cobertura y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Hyndman & Athanasopoulos, 2021; Moreno-Ponce et al., 2025).

La estimación utiliza validación de origen rodante y diagnóstico residual, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza

y la fecha en que se ejecutó el procedimiento (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

La interpretación se orienta a una decisión técnica definida: comprobar estabilidad del pronóstico a medida que avanza el tiempo. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En evaluación trimestral de modelos de seguridad, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Box et al., 2015; Feng et al., 2022).

Para visualizar esquema de validación temporal mediante origen rodante, la Figura 4.2 presenta una representación técnica que complementa la explicación y facilita su reproducción.

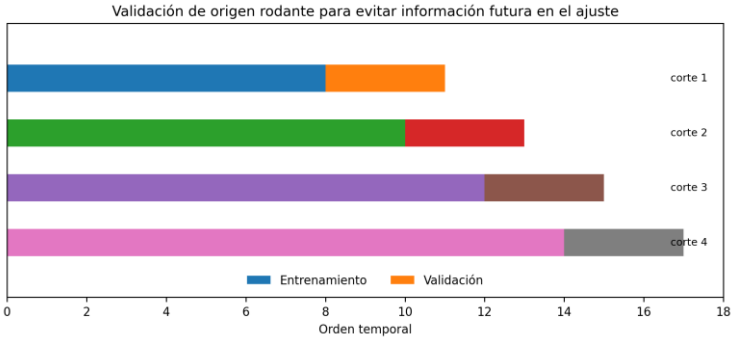


Figura 4.2. Esquema de validación temporal mediante origen rodante.

Nota. Elaboración propia. Cada corte conserva el orden cronológico y desplaza la ventana de evaluación hacia el futuro.

4.5.3 Error absoluto, error cuadrático y cobertura de intervalos

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean residuos, horizontes, errores y cobertura y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de

codificación, precisión o cobertura entre campañas de levantamiento (Instituto Nacional de Estadística y Censos, 2025; Feng et al., 2022).

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \qquad \mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$$

MAE y RMSE cuantifican el error con criterios diferentes. RMSE penaliza con mayor intensidad los errores grandes, mientras que MAE conserva una interpretación directa en las unidades de la variable pronosticada.

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando validación de origen rodante y diagnóstico residual como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El criterio de selección se define antes de observar el conjunto reservado para prueba (Moreno-Ponce et al., 2025; Instituto Nacional de Estadística y Censos, 2025).

La interpretación se orienta a una decisión técnica definida: comprobar estabilidad del pronóstico a medida que avanza el tiempo. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En evaluación trimestral de modelos de seguridad, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Moreno-Ponce et al., 2025; Box et al., 2015).

Para ordenar métricas para evaluación de pronósticos, la Tabla 4.2 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 4-2. Métricas para evaluación de pronósticos

Métrica	Ventaja	Precaución
MAE	Lectura directa del error	No penaliza fuertemente extremos

RMSE	Penaliza errores grandes	Sensible a valores extremos
MASE	Comparable entre series	Requiere línea base
Cobertura	Evalúa intervalos	Debe acompañarse de ancho
Sesgo medio	Detecta sobre/subpredicción	Puede compensar errores

Nota. Elaboración propia. Las métricas deben reportarse por horizonte cuando la decisión cambia con el tiempo.

4.5.4 Estabilidad del modelo frente a nuevos periodos observados

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran residuos, horizontes, errores y cobertura, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Moreno-Ponce et al., 2025; Feng et al., 2022).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora validación de origen rodante y diagnóstico residual. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El criterio de selección se define antes de observar el conjunto reservado para prueba (Hamilton, 1994; Moreno-Ponce et al., 2025).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: comprobar estabilidad del pronóstico a medida que avanza el tiempo. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En evaluación trimestral de modelos de seguridad, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción

impide presentar una predicción como si fuera evidencia de campo (Hamilton, 1994; Instituto Nacional de Estadística y Censos, 2025).

Ejercicio 4.5. Validación, diagnóstico y ajuste de ARIMA

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con residuos del modelo ARIMA y observaciones de prueba. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Revise media residual, autocorrelación, normalidad aproximada y estabilidad; calcule MAE y RMSE en el periodo reservado.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: diagnóstico de validez y decisión de mantener o reajustar el modelo. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

4.6 Implementación computacional de ARIMA con datos de siniestros

El análisis de implementación computacional de arima con datos de siniestros, desde una perspectiva de ingeniería civil, requiere vincular el cálculo con condiciones observables de la infraestructura. El análisis combina serie limpia, especificación, parámetros y pronósticos y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría,

la operación o la infraestructura que posteriormente deberá inspeccionarse (Hyndman & Athanasopoulos, 2021; Moreno-Ponce et al., 2025).

En la práctica, statsmodels y flujo reproducible en Python permite organizar la información para automatizar sin ocultar decisiones estadísticas. El análisis se plantea para laboratorios académicos y análisis institucionales, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Feng et al., 2022; Hyndman & Athanasopoulos, 2021).

4.6.1 Preparación reproducible de la serie en Python

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran serie limpia, especificación, parámetros y pronósticos, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Moreno-Ponce et al., 2025; Hamilton, 1994).

Para mantener la reproducibilidad, el análisis incorpora statsmodels y flujo reproducible en Python dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El criterio de selección se define antes de observar el conjunto reservado para prueba (Hamilton, 1994; Instituto Nacional de Estadística y Censos, 2025).

En términos de ingeniería, el análisis orienta la decisión de automatizar sin ocultar decisiones estadísticas. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En laboratorios académicos y análisis institucionales, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría

comprobar si la prioridad asignada tiene respaldo en terreno (Hamilton, 1994; Box et al., 2015).

4.6.2 Estimación de parámetros y registro de especificaciones

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran serie limpia, especificación, parámetros y pronósticos, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Moreno-Ponce et al., 2025; Box et al., 2015).

Para mantener la reproducibilidad, el análisis incorpora statsmodels y flujo reproducible en Python dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El criterio de selección se define antes de observar el conjunto reservado para prueba (Hamilton, 1994; Feng et al., 2022).

4.6.3 Pronóstico, intervalos y comparación con observaciones

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran serie limpia, especificación, parámetros y pronósticos, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Box et al., 2015; Feng et al., 2022).

La estimación utiliza statsmodels y flujo reproducible en Python, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Una

partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Instituto Nacional de Estadística y Censos, 2025; Hamilton, 1994).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: automatizar sin ocultar decisiones estadísticas. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En laboratorios académicos y análisis institucionales, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Instituto Nacional de Estadística y Censos, 2025; Box et al., 2015).

La evaluación temporal debe mostrar simultáneamente la trayectoria observada, el pronóstico y su incertidumbre. La Figura 4.3 presenta un ejemplo SARIMA ajustado sobre la serie trimestral oficial utilizada en el capítulo y extiende la estimación hacia 2026.

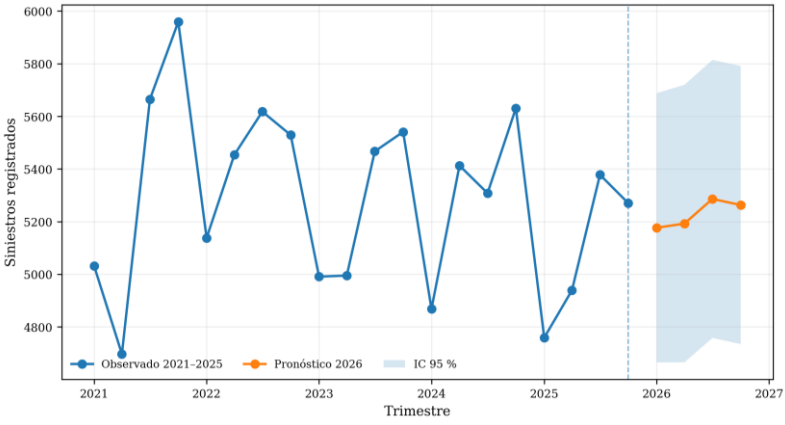


Figura 4.3. Pronóstico SARIMA trimestral de siniestros con intervalo de confianza.

Nota. Elaboración propia a partir de datos INEC–ANT, 2021–2025; el tramo 2026 corresponde a estimaciones del modelo.

El intervalo de confianza hace visible la incertidumbre creciente del pronóstico y evita presentar una única línea futura como resultado determinista. Su utilización profesional requiere actualización cuando se incorporen nuevos registros y comparación frente a una línea base más simple.

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Hyndman & Athanasopoulos, 2021; Feng et al., 2022).

4.6.4 Actualización del modelo cuando se incorpora nueva información

La evaluación se sustenta en serie limpia, especificación, parámetros y pronósticos, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Feng et al., 2022; Instituto Nacional de Estadística y Censos, 2025).

La estimación utiliza statsmodels y flujo reproducible en Python, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Hyndman & Athanasopoulos, 2021; Feng et al., 2022).

Ejercicio 4.6. Implementación computacional de ARIMA

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con serie mensual de siniestros con fecha y conteo. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Implemente ARIMA con statsmodels, produzca pronóstico de 12 meses e intervalos de confianza y exporte los resultados a CSV.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: tabla de pronóstico y gráfico observado-pronosticado. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Ejemplo breve. Ejemplo de partición: con 84 meses, use los primeros 72 para ajuste y los últimos 12 para prueba. MAE y RMSE deben calcularse únicamente sobre esos 12 meses.

4.7 Extensiones del modelo: SARIMA y ARIMAX con variables exógenas

La comprensión de extensiones del modelo: sarima y arimax con variables exógenas requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran estacionalidad, lluvia, movilidad, feriados y cambios operativos; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Moreno-Ponce et al., 2025; Box et al., 2015).

El procedimiento adoptado emplea SARIMA, ARIMAX e intervención y vincula cada salida con una acción verificable: incorporar información exógena cuando existe soporte temporal. La aplicación en pronósticos con variables explicativas disponibles antes del periodo a predecir obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (Instituto Nacional de Estadística y Censos, 2025; Feng et al., 2022).

4.7.1 Estacionalidad explícita mediante SARIMA

Para abordar el problema se integran estacionalidad, lluvia, movilidad, feriados y cambios operativos, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Feng et al., 2022; Box et al., 2015).

$$\Phi_P(B^s)\Phi_p(B)(1-B)^d(1-B^s)^D y_t = \Theta_Q(B^s)\Theta_q(B)\varepsilon_t$$

SARIMA incorpora operadores estacionales con periodo s además de los componentes no estacionales. Esta estructura resulta útil cuando la siniestralidad presenta regularidades mensuales, trimestrales u otras repeticiones vinculadas al calendario.

La estimación utiliza SARIMA, ARIMAX e intervención, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Hyndman & Athanasopoulos, 2021; Feng et al., 2022).

Para la ingeniería, el análisis resulta pertinente cuando permite incorporar información exógena cuando existe soporte temporal. Para ello se acompañan

la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En pronósticos con variables explicativas disponibles antes del periodo a predecir, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Hyndman & Athanasopoulos, 2021; Hamilton, 1994).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Hamilton, 1994; Instituto Nacional de Estadística y Censos, 2025).

4.7.2 Variables exógenas y formulación de ARIMAX

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean estacionalidad, lluvia, movilidad, feriados y cambios operativos y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Hamilton, 1994; Instituto Nacional de Estadística y Censos, 2025).

$$y_t = \beta_0 + \boldsymbol{\beta}^T \mathbf{x}_t + n_t$$

En ARIMAX, el término β^T x_t representa el aporte de variables externas, mientras que n_t conserva una estructura ARIMA. Las variables exógenas deben estar disponibles en el momento real de efectuar el pronóstico.

La validación exige contrastar la salida obtenida mediante SARIMA, ARIMAX e intervención con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización,

no solo en el valor obtenido durante el entrenamiento (Feng et al., 2022; Box et al., 2015).

En términos de ingeniería, el análisis orienta la decisión de incorporar información exógena cuando existe soporte temporal. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En pronósticos con variables explicativas disponibles antes del periodo a predecir, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Feng et al., 2022; Hamilton, 1994).

Una comprobación adicional consiste en retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Moreno-Ponce et al., 2025; Instituto Nacional de Estadística y Censos, 2025).

4.7.3 Intervenciones, cambios normativos y variables indicadoras

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean estacionalidad, lluvia, movilidad, feriados y cambios operativos y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Instituto Nacional de Estadística y Censos, 2025; Hamilton, 1994).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora SARIMA, ARIMAX e intervención. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. El criterio de selección se define antes

de observar el conjunto reservado para prueba (Moreno-Ponce et al., 2025; Instituto Nacional de Estadística y Censos, 2025).

La aplicación del procedimiento cobra sentido al apoyar la decisión de incorporar información exógena cuando existe soporte temporal. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En pronósticos con variables explicativas disponibles antes del periodo a predecir, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Moreno-Ponce et al., 2025; Box et al., 2015).

4.7.4 Comparación entre modelos estacionales y no estacionales

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean estacionalidad, lluvia, movilidad, feriados y cambios operativos y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Hyndman & Athanasopoulos, 2021; Moreno-Ponce et al., 2025).

La estimación utiliza SARIMA, ARIMAX e intervención, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Una partición aleatoria rompe el orden temporal y puede incorporar información futura en el entrenamiento; la evaluación debe avanzar mediante cortes cronológicos sucesivos. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: incorporar información exógena cuando existe soporte temporal. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En pronósticos con variables explicativas disponibles antes del periodo a predecir, la recomendación conserva una separación estricta entre dato observado,

estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Box et al., 2015; Feng et al., 2022).

Ejercicio 4.7. Extensiones SARIMA y ARIMAX con variables exógenas

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con siniestros mensuales, lluvia, TPDA u otra variable externa y periodicidad anual. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Ajuste SARIMA y SARIMAX; compruebe que la variable externa esté disponible para el horizonte futuro y compare errores fuera de muestra.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: selección entre modelo estacional puro y modelo con variable exógena. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Laboratorio Python 4

El siguiente laboratorio implementa sarima con datos trimestrales de ecuador. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 4. SARIMA con siniestros trimestrales de Ecuador 2021-2025
# Requisitos: pandas, numpy, statsmodels, matplotlib
import numpy as np
import pandas as pd
import matplotlib. pyplot as plt
from statsmodels. tsa. statespace. sarimax import SARIMAX
from sklearn. metrics import mean_absolute_error, mean_squared_error

valores = [5031,4697,5665,5959, 5138,5454,5618,5529,
 4991,4995,5467,5541, 4868,5413,5308,5631,
 4759,4939,5378,5270]
idx = pd. period_range("2021Q1", periods=len(valores), freq="Q")
y = pd. Series(valores, index=idx, dtype=float)

# Validación: últimos cuatro trimestres como prueba.
train, test = y. iloc[:-4], y. iloc[-4:]
model = SARIMAX(
 train,
 order=(1,0,0),
 seasonal_order=(1,0,0,4),
 trend="c",
 enforce_stationarity=False,
 enforce_invertibility=False
). fit(disp=False)
forecast = model. get_forecast(steps=4)
pred = forecast. predicted_mean

mae = mean_absolute_error(test, pred)
rmse = mean_squared_error(test, pred) ** 0.5
print(f"MAE={mae:.1f}; RMSE={rmse:.1f}")
print(pd. DataFrame({"real": test, "pred": pred}))

# Reajuste con toda la serie y pronóstico de cuatro trimestres.
final = SARIMAX(
 y, order=(1,0,0), seasonal_order=(1,0,0,4), trend="c",
 enforce_stationarity=False, enforce_invertibility=False
). fit(disp=False)
future = final. get_forecast(steps=4)
print(future. summary_frame())

plt. figure(figsize=(8,4))
plt. plot(y. to_timestamp(), y, marker="o", label="Observado")
plt. plot(future. predicted_mean. to_timestamp(), future. predicted_mean,
```

```
 marker="o", label="Pronóstico")
ci = future. conf_int()
plt. fill_between(ci. index. to_timestamp(), ci. iloc[:,0], ci. iloc[:,1], alpha=.2)
plt. legend(); plt. title("SARIMA trimestral"); plt. tight_layout(); plt. show()
```

Preguntas para el laboratorio

• ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?

• ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?

• ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?

• ¿Qué información de campo agregaría para reducir incertidumbre?

• ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

4.8 Referencias del capítulo 4

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley.

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts.

Feng, T., Zheng, Z., Xu, J., Liu, M., Li, M., Jia, H., & Yu, X. (2022). The comparative analysis of SARIMA, Facebook Prophet, and LSTM for road traffic injury prediction in Northeast China. Frontiers in Public Health, 10, 946563. https://doi. org/10.3389/fpubh.2022.946563

Hamilton, J. D. (1994). Time series analysis. Princeton University Press.

Moreno-Ponce, L. A., Pérez-Zuriaga, A. M., & García, A. (2025). Predictive models and GIS for road safety: Application to a segment of the Chone–Flavio Alfaro road. Sustainability, 17(11), 5032. https://doi. org/10.3390/su17115032

Instituto Nacional de Estadística y Censos. (2024). Estadísticas de Transporte 2023. INEC. https://www. ecuadorencifras. gob. ec/estadisticas-de-transporte-2023/

Instituto Nacional de Estadística y Censos. (2025). Siniestros de Tránsito IV trimestre 2024. INEC. https://www. ecuadorencifras. gob. ec/siniestros-de-transito-iv-trimestre-2024/

Instituto Nacional de Estadística y Censos. (2026). Siniestros de Tránsito IV trimestre 2025. INEC. https://www. ecuadorencifras. gob. ec/siniestros-de-transito-iv-trimestre-2025/

Skaug, L., Nojoumian, M., Dang, N., & Yap, A. (2025). Road crash analysis and modeling: A systematic review of methods, data, and emerging technologies. Applied Sciences, 15(13), 7115. https://doi. org/10.3390/app15137115

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO V

Redes Neuronales LSTM para Predicción Espacio-Temporal

AUTOR: González Regalado Marlon Alexander

EDITORIAL RUNAIKI

5 Redes Neuronales LSTM para Predicción Espacio-Temporal

Las redes LSTM son útiles cuando la secuencia presenta relaciones no lineales y memoria de varios pasos. Su ventaja no es automática: requieren suficientes datos, separación temporal estricta, normalización y comparación con modelos simples. Este capítulo emplea PyTorch para mantener un laboratorio ejecutable sin depender de servicios externos. Las revisiones de aprendizaje automático aplicado a siniestralidad insisten en la calidad del dato, el desbalance, la transferencia espacial y el control de fuga temporal antes de interpretar una mejora de desempeño (Chai et al., 2024; Butt & Shafique, 2025).

Resultados de aprendizaje

- Describir el funcionamiento de una LSTM y su diferencia frente a una red recurrente convencional.
- Construir ventanas temporales sin fuga de información y con escalamiento reproducible.
- Entrenar redes con regularización, parada temprana y particiones cronológicas.
- Comparar LSTM con líneas base estadísticas usando métricas coherentes con el horizonte.
- Definir condiciones de despliegue, monitoreo y reentrenamiento para un sistema de predicción.

5.1 Fundamentos de redes neuronales recurrentes en transporte

Desde la ingeniería civil, el estudio de fundamentos de redes neuronales recurrentes en transporte exige algo más que ejecutar una rutina estadística. El análisis combina secuencias, estados ocultos y dependencias largas y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Feng et al., 2022; Butt & Shafique, 2025).

Desde una perspectiva operativa, RNN y LSTM ayuda a estructurar la información y a capturar memoria temporal cuando las relaciones no son bien representadas por modelos lineales. El análisis se plantea para series viales con información multivariable, donde la continuidad de los registros puede ser más

importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Chai et al., 2024; Goodfellow et al., 2016).

5.1.1 Memoria secuencial en redes neuronales recurrentes

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de secuencias, estados ocultos y dependencias largas se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Feng et al., 2022; Butt & Shafique, 2025).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora RNN y LSTM. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chollet, 2021; Goodfellow et al., 2016).

La aplicación del procedimiento cobra sentido al apoyar la decisión de capturar memoria temporal cuando las relaciones no son bien representadas por modelos lineales. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En series viales con información multivariable, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Chollet, 2021; Feng et al., 2022).

5.1.2 Gradientes inestables y motivación de las unidades LSTM

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran secuencias, estados ocultos y dependencias largas, diferenciando los atributos

que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chollet, 2021; Chai et al., 2024).

La estimación se obtiene mediante RNN y LSTM y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Goodfellow et al., 2016; Butt & Shafique, 2025).

5.1.3 Longitud de secuencia y dependencia temporal útil

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean secuencias, estados ocultos y dependencias largas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Goodfellow et al., 2016; Chai et al., 2024).

La estimación utiliza RNN y LSTM, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Hochreiter & Schmidhuber, 1997; Goodfellow et al., 2016).

Los resultados deben interpretarse con una finalidad explícita: capturar memoria temporal cuando las relaciones no son bien representadas por

modelos lineales. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series viales con información multivariable, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

5.1.4 Aplicaciones de RNN en series de movilidad y seguridad vial

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran secuencias, estados ocultos y dependencias largas, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chollet, 2021; Chai et al., 2024).

La evaluación utiliza RNN y LSTM, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Goodfellow et al., 2016; Butt & Shafique, 2025).

Los resultados deben interpretarse con una finalidad explícita: capturar memoria temporal cuando las relaciones no son bien representadas por modelos lineales. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series viales con información multivariable, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Goodfellow et al., 2016; Chollet, 2021).

Ejercicio 5.1. Redes neuronales recurrentes aplicadas al transporte

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con secuencia mensual de siniestros y ventana de 12 meses. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Forme pares entrada-salida, escale solo con el conjunto de entrenamiento y construya una red recurrente sencilla como línea base.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: matrices X-y con dimensiones verificadas y primera predicción. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

5.2 Arquitectura LSTM y modelación de dependencias temporales

En arquitectura lstm y modelación de dependencias temporales, la elección del algoritmo ocupa un segundo plano frente a la organización y calidad de la evidencia disponible. La base reúne estado de celda, puertas y activaciones y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Chollet, 2021; Feng et al., 2022).

La aplicación práctica de la arquitectura LSTM permite disponer la información de manera que sea posible entender cómo se conserva o descarta información en cada paso temporal. El análisis se plantea para modelos de

predicción de siniestralidad y movilidad, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Feng et al., 2022; Butt & Shafique, 2025).

5.2.1 Estado de celda y circulación controlada de información

Para abordar el problema se integran estado de celda, puertas y activaciones, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hochreiter & Schmidhuber, 1997; Goodfellow et al., 2016).

Para sostener la interpretación, el procedimiento basado en la arquitectura LSTM se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La transformación de escala se calibra con el conjunto de entrenamiento y se conserva sin cambios para los periodos de validación y prueba. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Butt & Shafique, 2025; Goodfellow et al., 2016).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: entender cómo se conserva o descarta información en cada paso temporal. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En modelos de predicción de siniestralidad y movilidad, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Butt & Shafique, 2025; Feng et al., 2022).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible

para priorizar recursos de inspección (Goodfellow et al., 2016; Butt & Shafique, 2025).

5.2.2 Puertas de olvido, entrada y salida

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean estado de celda, puertas y activaciones y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Feng et al., 2022; Goodfellow et al., 2016).

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \qquad h_t = o_t \odot \tanh(C_t)$$

Las puertas regulan qué información se conserva, incorpora o transmite. El operador $\odot$ representa el producto elemento a elemento y σ la función sigmoide; esta arquitectura permite conservar dependencias temporales durante secuencias más largas.

La validación exige contrastar la salida obtenida mediante la arquitectura LSTM con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Chollet, 2021; Feng et al., 2022).

En términos de ingeniería, el análisis orienta la decisión de entender cómo se conserva o descarta información en cada paso temporal. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En modelos de predicción de siniestralidad y movilidad, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Chollet, 2021; Chai et al., 2024).

Para visualizar arquitectura funcional simplificada de una celda lstm, la Figura 5.1 presenta una representación técnica que complementa la explicación y facilita su reproducción.

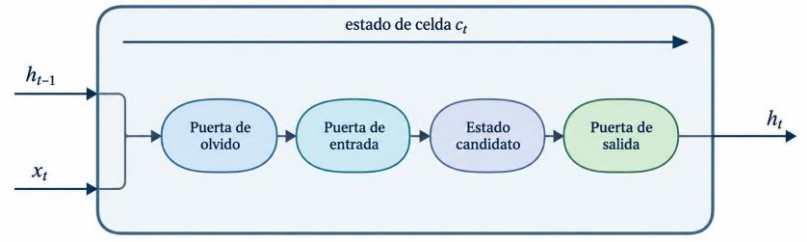


Figura 5.1. Arquitectura funcional simplificada de una celda LSTM.

Nota. Elaboración propia con base en Hochreiter y Schmidhuber (1997).

5.2.3 Dimensiones de tensores en entradas multivariables

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de estado de celda, puertas y activaciones se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Feng et al., 2022; Chai et al., 2024).

Para mantener la reproducibilidad, el análisis incorpora la arquitectura LSTM dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chollet, 2021; Butt & Shafique, 2025).

5.2.4 Interpretación funcional de activaciones y estados ocultos

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean estado de celda, puertas y activaciones y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o

cobertura entre campañas de levantamiento (Goodfellow et al., 2016; Chai et al., 2024).

La estimación utiliza arquitectura LSTM, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. La transformación de escala se calibra con el conjunto de entrenamiento y se conserva sin cambios para los periodos de validación y prueba. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Hochreiter & Schmidhuber, 1997; Goodfellow et al., 2016).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: entender cómo se conserva o descarta información en cada paso temporal. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En modelos de predicción de siniestralidad y movilidad, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

Ejercicio 5.2. Arquitectura LSTM y dependencias temporales

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con una secuencia normalizada y pesos didácticos de puertas LSTM. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule para un paso las puertas de olvido, entrada y salida; actualice estado de celda y estado oculto.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: comprobación numérica de un paso antes de usar una biblioteca de aprendizaje profundo. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

5.3 Preprocesamiento de datos multivariables espaciales y temporales

En preprocesamiento de datos multivariables espaciales y temporales, la primera decisión consiste en precisar qué representa cada observación y qué pregunta deberá resolver el análisis. Se consideran escalas, rezagos, ventanas y particiones cronológicas, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Butt & Shafique, 2025; Hochreiter & Schmidhuber, 1997).

La sección articula transformación supervisada de secuencias con una finalidad profesional concreta: evitar fuga de información durante el preprocesamiento. En series con TPDA, clima y variables de operación, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

5.3.1 Escalamiento robusto de variables con unidades diferentes

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran escalas, rezagos, ventanas y particiones cronológicas, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra

la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chollet, 2021; Feng et al., 2022).

$$x_i^* = \frac{x_i - \text{mediana}(x)}{\text{IQR}(x)}$$

El escalamiento robusto utiliza la mediana y el rango intercuartílico, por lo que reduce la influencia de valores extremos. Sus parámetros deben estimarse únicamente con los datos de entrenamiento para evitar fuga de información.

El procedimiento combina la transformación supervisada de secuencias con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. El escalamiento debe aprenderse exclusivamente con los datos de entrenamiento. Los mismos parámetros se transfieren después a validación y prueba para evitar fuga de información. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Goodfellow et al., 2016; Chai et al., 2024).

La interpretación técnica orienta una acción verificable: evitar fuga de información durante el preprocesamiento. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En series con TPDA, clima y variables de operación, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Goodfellow et al., 2016; Hochreiter & Schmidhuber, 1997).

5.3.2 Construcción de rezagos y ventanas supervisadas

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean escalas, rezagos, ventanas y particiones cronológicas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Goodfellow et al., 2016; Feng et al., 2022).

La estimación utiliza transformación supervisada de secuencias, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. El escalamiento debe aprenderse exclusivamente con los datos de entrenamiento. Los mismos parámetros se transfieren después a validación y prueba para evitar fuga de información. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Hochreiter & Schmidhuber, 1997; Butt & Shafique, 2025).

La interpretación técnica orienta una acción verificable: evitar fuga de información durante el preprocesamiento. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En series con TPDA, clima y variables de operación, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Hochreiter & Schmidhuber, 1997; Goodfellow et al., 2016).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Chollet, 2021; Feng et al., 2022).

5.3.3 Separación cronológica antes de transformar los datos

Para abordar el problema se integran escalas, rezagos, ventanas y particiones cronológicas, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Chai et al., 2024; Chollet, 2021).

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante la transformación supervisada de secuencias frente a una referencia

ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. La transformación de escala se calibra con el conjunto de entrenamiento y se conserva sin cambios para los periodos de validación y prueba. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Feng et al., 2022; Chai et al., 2024).

Para la ingeniería, el análisis resulta pertinente cuando permite evitar fuga de información durante el preprocesamiento. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En series con TPDA, clima y variables de operación, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Feng et al., 2022; Butt & Shafique, 2025).

Para ordenar preparación de entradas para una lstm, la Tabla 5.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 5-1. Preparación de entradas para una LSTM

Etapa	Regla	Riesgo evitado
Ordenar	Secuencia cronológica	Desalineación
Particionar	Antes de escalar	Fuga de información
Escalar	Ajustar solo con entrenamiento	Uso de datos futuros
Ventanas	Definir longitud fija	Inconsistencia temporal
Objetivo	Separar horizonte	Predicción ambigua

Nota. Elaboración propia con base en Goodfellow et al. (2016) y Chollet (2021).

5.3.4 Tratamiento de faltantes sin contaminar periodos futuros

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean escalas, rezagos, ventanas y particiones cronológicas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o

transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Butt & Shafique, 2025; Goodfellow et al., 2016).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando la transformación supervisada de secuencias como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. El escalamiento debe aprenderse exclusivamente con los datos de entrenamiento. Los mismos parámetros se transfieren después a validación y prueba para evitar fuga de información. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chai et al., 2024; Chollet, 2021).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de evitar fuga de información durante el preprocesamiento. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En series con TPDA, clima y variables de operación, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Chai et al., 2024; Butt & Shafique, 2025).

Ejercicio 5.3. Preprocesamiento de datos multivariables

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con siniestros, lluvia, V85, TPDA y severidad por periodo. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Alinee fechas, trate faltantes, escale cada variable y transforme el arreglo a formato muestras-tiempo-variables.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: tensor sin fuga de información y registro de parámetros de escalado. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

5.4 Entrenamiento, validación y ajuste de modelos LSTM

Desde la ingeniería civil, el estudio de entrenamiento, validación y ajuste de modelos lstm exige algo más que ejecutar una rutina estadística. El análisis combina unidades, capas, lotes, dropout y épocas y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Feng et al., 2022; Hochreiter & Schmidhuber, 1997).

Desde una perspectiva operativa, el entrenamiento con parada temprana ayuda a estructurar la información y a controlar sobreajuste y reproducibilidad. El análisis se plantea para muestras temporales de tamaño limitado, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Chai et al., 2024; Feng et al., 2022).

5.4.1 Selección de unidades, capas y tamaño de lote

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean unidades, capas, lotes, dropout y épocas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Butt & Shafique, 2025; Feng et al., 2022).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora el entrenamiento con parada temprana. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chai et al., 2024; Butt & Shafique, 2025).

5.4.2 Regularización, dropout y control del sobreajuste

En este apartado, la evidencia se organiza a partir de unidades, capas, lotes, dropout y épocas. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Chai et al., 2024; Chollet, 2021).

La validación exige contrastar la salida obtenida mediante el entrenamiento con parada temprana con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Feng et al., 2022; Chai et al., 2024).

El resultado es útil cuando permite controlar sobreajuste y reproducibilidad. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En muestras temporales de tamaño limitado, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Feng et al., 2022; Butt & Shafique, 2025).

Para visualizar curvas ilustrativas de entrenamiento y validación para control de sobreajuste, la Figura 5.2 presenta una representación técnica que complementa la explicación y facilita su reproducción.

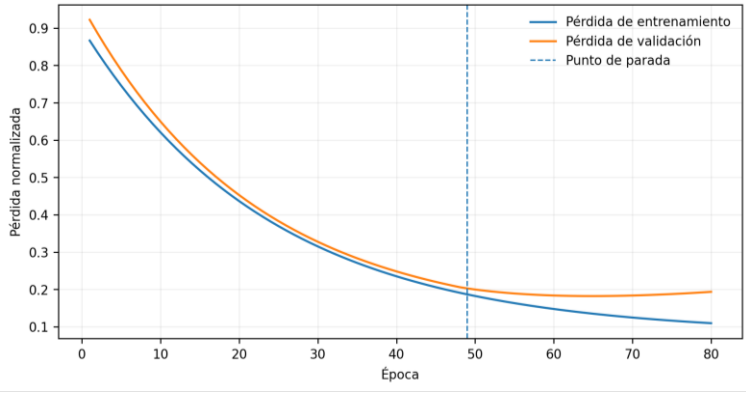


Figura 5.2. Curvas ilustrativas de entrenamiento y validación para control de sobreajuste.

Nota. Elaboración propia. Las curvas son didácticas y permiten explicar la parada temprana.

5.4.3 Parada temprana basada en validación temporal

En este apartado, la evidencia se organiza a partir de unidades, capas, lotes, dropout y épocas. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hochreiter & Schmidhuber, 1997; Feng et al., 2022).

La evaluación utiliza el entrenamiento con parada temprana, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Butt & Shafique, 2025; Hochreiter & Schmidhuber, 1997).

En términos de ingeniería, el análisis orienta la decisión de controlar sobreajuste y reproducibilidad. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para

que otro equipo pueda repetir el cálculo. En muestras temporales de tamaño limitado, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Butt & Shafique, 2025; Goodfellow et al., 2016).

5.4.4 Repetibilidad del entrenamiento y manejo de semillas aleatorias

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran unidades, capas, lotes, dropout y épocas, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chollet, 2021; Butt & Shafique, 2025).

La evaluación utiliza el entrenamiento con parada temprana, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Goodfellow et al., 2016; Chollet, 2021).

La lectura ingenieril del resultado debe traducirse en una acción concreta: controlar sobreajuste y reproducibilidad. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En muestras temporales de tamaño limitado, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Goodfellow et al., 2016; Feng et al., 2022).

Ejercicio 5.4. Entrenamiento, validación y ajuste de LSTM

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con tensores de entrenamiento, validación y prueba. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Entrene con early stopping, guarde el mejor estado por pérdida de validación y evite usar la prueba para escoger épocas o hiperparámetros.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: curvas de pérdida y modelo retenido en la mejor época. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

5.5 Comparación cuantitativa entre ARIMA y LSTM

La comprensión de comparación cuantitativa entre arima y lstm requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran MAE, RMSE, horizonte y costo computacional; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Chai et al., 2024; Hochreiter & Schmidhuber, 1997).

La aplicación práctica de la comparación ARIMA-LSTM permite disponer la información de manera que sea posible demostrar una ganancia real frente a una línea base robusta. El análisis se plantea para series de seguridad vial con distinta longitud, donde la continuidad de los registros puede ser más

importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Butt & Shafique, 2025; Chollet, 2021).

5.5.1 Línea base estadística antes de evaluar una LSTM

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de MAE, RMSE, horizonte y costo computacional se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Goodfellow et al., 2016; Hochreiter & Schmidhuber, 1997).

Para sostener la interpretación, el procedimiento basado en la comparación ARIMA-LSTM se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La transformación de escala se calibra con el conjunto de entrenamiento y se conserva sin cambios para los periodos de validación y prueba. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: demostrar una ganancia real frente a una línea base robusta. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series de seguridad vial con distinta longitud, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Hochreiter & Schmidhuber, 1997; Feng et al., 2022).

5.5.2 Comparación por horizontes y no solo por promedio global

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean MAE, RMSE, horizonte y costo computacional y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y

temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Butt & Shafique, 2025; Feng et al., 2022).

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$$

El error cuadrático medio resume la magnitud de las diferencias entre observaciones y predicciones. En series viales conviene calcularlo también por horizonte, porque un promedio global puede ocultar deterioros importantes en predicciones más lejanas.

La aplicación metodológica de la comparación ARIMA-LSTM sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. La transformación de escala se calibra con el conjunto de entrenamiento y se conserva sin cambios para los periodos de validación y prueba. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chai et al., 2024; Butt & Shafique, 2025).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de demostrar una ganancia real frente a una línea base robusta. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En series de seguridad vial con distinta longitud, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Chai et al., 2024; Hochreiter & Schmidhuber, 1997).

Para que la comparación entre ARIMA y LSTM sea comprensible, conviene representar el comportamiento del modelo sobre observaciones reservadas que no intervinieron en el ajuste. La Figura 5.3 ilustra la predicción de una LSTM sobre una serie mensual derivada con fines docentes.

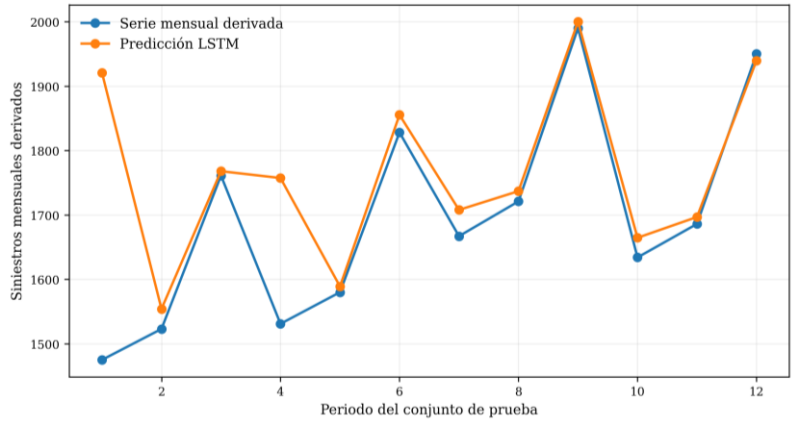


Figura 5.3. Predicción LSTM frente a valores derivados reservados para prueba.

Nota. Elaboración propia a partir del Laboratorio Python 5. La serie mensual es derivada de totales trimestrales y no constituye un registro mensual oficial.

La cercanía entre ambas curvas debe evaluarse con métricas fuera de muestra y no únicamente por inspección visual. Las desviaciones persistentes permiten detectar retrasos, sobreajuste o incapacidad del modelo para reproducir cambios rápidos, aspectos que condicionan su uso operativo.

Una comprobación adicional consiste en retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

5.5.3 Tamaño muestral frente a capacidad del modelo

La evaluación se sustenta en MAE, RMSE, horizonte y costo computacional, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha

y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

La estimación se obtiene mediante la comparación ARIMA-LSTM y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Butt & Shafique, 2025; Chai et al., 2024).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: demostrar una ganancia real frente a una línea base robusta. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series de seguridad vial con distinta longitud, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Butt & Shafique, 2025; Hochreiter & Schmidhuber, 1997).

Para ordenar comparación técnica entre arima y lstm, la Tabla 5.2 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 5-2. Comparación técnica entre ARIMA y LSTM

Criterio	ARIMA	LSTM
Datos requeridos	Bajo-medio	Medio-alto
Interpretabilidad	Alta	Media-baja
No linealidad	Limitada	Alta capacidad
Entrenamiento	Rápido	Mayor costo
Validación temporal	Obligatoria	Obligatoria

Nota. Elaboración propia. La elección depende del desempeño fuera de muestra y de la capacidad institucional para mantener el modelo.

5.5.4 Costo computacional y utilidad marginal de la red neuronal

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean MAE, RMSE, horizonte y costo computacional y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Feng et al., 2022; Hochreiter & Schmidhuber, 1997).

La validación exige contrastar la salida obtenida mediante la comparación ARIMA-LSTM con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Chollet, 2021; Feng et al., 2022).

Los resultados deben interpretarse con una finalidad explícita: demostrar una ganancia real frente a una línea base robusta. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En series de seguridad vial con distinta longitud, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Chollet, 2021; Chai et al., 2024).

Ejercicio guiado 5.5. Comparación cuantitativa entre ARIMA y LSTM

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con pronósticos ARIMA y LSTM sobre los mismos 12 meses de prueba. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.

2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule MAE, RMSE y error por mes; aplique una prueba de sensibilidad retirando el mes con mayor error.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: comparación justa bajo el mismo horizonte y partición. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

5.6 Modelos híbridos ARIMA-LSTM y CNN-LSTM

La comprensión de modelos híbridos arima-lstm y cnn-lstm requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran residuos, filtros temporales y pronósticos combinados; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Hochreiter & Schmidhuber, 1997; Chai et al., 2024).

En la práctica, ARIMA-LSTM, CNN-LSTM y ensamblados permite organizar la información para aprovechar componentes complementarios sin añadir complejidad innecesaria. El análisis se plantea para series con no linealidad y estacionalidad, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Goodfellow et al., 2016; Chollet, 2021).

5.6.1 Híbridos basados en residuos ARIMA-LSTM

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran residuos, filtros temporales y pronósticos combinados, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Hochreiter & Schmidhuber, 1997; Goodfellow et al., 2016).

La estimación utiliza ARIMA-LSTM, CNN-LSTM y ensamblados, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Butt & Shafique, 2025; Chollet, 2021).

La lectura ingenieril del resultado debe traducirse en una acción concreta: aprovechar componentes complementarios sin añadir complejidad innecesaria. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En series con no linealidad y estacionalidad, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Butt & Shafique, 2025; Chai et al., 2024).

5.6.2 Extracción local de características mediante CNN temporal

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de residuos, filtros temporales y pronósticos combinados se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si

representan un error o una condición vial real que precisamente interesa estudiar (Goodfellow et al., 2016; Butt & Shafique, 2025).

La evaluación utiliza ARIMA-LSTM, CNN-LSTM y ensamblados, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Hochreiter & Schmidhuber, 1997; Goodfellow et al., 2016).

La lectura ingenieril del resultado debe traducirse en una acción concreta: aprovechar componentes complementarios sin añadir complejidad innecesaria. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En series con no linealidad y estacionalidad, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

5.6.3 Combinación de pronósticos y ensamblados

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de residuos, filtros temporales y pronósticos combinados se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Feng et al., 2022; Chai et al., 2024).

Para mantener la reproducibilidad, el análisis incorpora ARIMA-LSTM, CNN-LSTM y ensamblados dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. Utilizar toda la serie para estimar el escalamiento contamina la evaluación. Por ello, la transformación se ajusta en entrenamiento y se aplica posteriormente a los demás conjuntos. El criterio de

selección se define antes de observar el conjunto reservado para prueba (Chollet, 2021; Chai et al., 2024).

5.6.4 Condiciones en las que un modelo híbrido no aporta mejora

Para abordar el problema se integran residuos, filtros temporales y pronósticos combinados, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

Para sostener la interpretación, el procedimiento basado en ARIMA-LSTM, CNN-LSTM y ensamblados se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La transformación de escala se calibra con el conjunto de entrenamiento y se conserva sin cambios para los periodos de validación y prueba. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Butt & Shafique, 2025; Feng et al., 2022).

Ejercicio 5.6. Modelos híbridos ARIMA-LSTM y CNN-LSTM

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con pronóstico ARIMA, residuos y variables adicionales. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Ajuste ARIMA al componente lineal, entrene LSTM con los residuos y sume ambas salidas; compare frente a cada modelo individual.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: verificación de si la combinación reduce error fuera de muestra. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

5.7 Implementación de sistemas de predicción en tiempo real

El análisis de implementación de sistemas de predicción en tiempo real parte de una definición explícita de la unidad de análisis y de la pregunta técnica que se pretende responder. Se consideran latencia, deriva, versiones y alertas, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Feng et al., 2022; Chai et al., 2024).

La incorporación del modelo a sistemas institucionales de alerta vial exige un despliegue controlado, monitoreo documentado, datos versionados y reglas de decisión definidas con anticipación. La prioridad consiste en mantener desempeño y supervisión humana. El desempeño estadístico se valora únicamente cuando mejora esa decisión. El resultado adquiere valor práctico cuando puede contrastarse con evidencia independiente, ya provenga de inspección, normativa, medición de campo o un periodo posterior (Chai et al., 2024; Hochreiter & Schmidhuber, 1997).

5.7.1 Ingesta de datos y latencia admisible del sistema

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran latencia, deriva, versiones y alertas, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede

reconstruirse sin depender de decisiones implícitas del analista (Chollet, 2021; Chai et al., 2024).

La estimación se obtiene mediante el despliegue y monitoreo de modelos y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Goodfellow et al., 2016; Butt & Shafique, 2025).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: mantener desempeño y supervisión humana. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En sistemas institucionales de alerta vial, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Goodfellow et al., 2016; Chollet, 2021).

5.7.2 Detección de deriva y pérdida de desempeño

La evaluación se sustenta en latencia, deriva, versiones y alertas, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Chai et al., 2024; Chollet, 2021).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante el despliegue y monitoreo de modelos se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Feng et al., 2022; Chai et al., 2024).

La utilidad del análisis se hace evidente cuando permite mantener desempeño y supervisión humana. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En sistemas institucionales de alerta vial, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Feng et al., 2022; Butt & Shafique, 2025).

5.7.3 Reentrenamiento, versionado y trazabilidad operativa

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean latencia, deriva, versiones y alertas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Goodfellow et al., 2016; Chai et al., 2024).

La estimación utiliza despliegue y monitoreo de modelos, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Hochreiter & Schmidhuber, 1997; Goodfellow et al., 2016).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: mantener desempeño y supervisión humana. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En sistemas institucionales de alerta vial, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

5.7.4 Supervisión humana de alertas antes de intervenir la vía

La evaluación se sustenta en latencia, deriva, versiones y alertas, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Hochreiter & Schmidhuber, 1997; Chollet, 2021).

La estimación se obtiene mediante el despliegue y monitoreo de modelos y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. Los parámetros de escalamiento se estiman solo con el conjunto de entrenamiento y luego se aplican a validación y prueba; calcularlos con toda la serie introduciría información futura. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Butt & Shafique, 2025; Chai et al., 2024).

Ejercicio 5.7. Sistemas de predicción en tiempo real

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con flujo de registros nuevos, modelo entrenado, fecha de actualización y umbral de alerta. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Diseñe una rutina que valide datos entrantes, calcule predicción, registre versión del modelo y emita alerta solo cuando se cumplan dos criterios.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: diagrama operativo y registro auditable de inferencias. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.

6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Laboratorio Python 5

El siguiente laboratorio implementa lstm con pytorch. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 5. LSTM con PyTorch sobre una serie mensual derivada
# Requisitos: numpy, torch, matplotlib
# La serie mensual se deriva de totales trimestrales oficiales con pesos fijos.
# No debe interpretarse como observación mensual publicada por INEC/ANT.
import numpy as np
import torch
import torch. nn as nn
import matplotlib. pyplot as plt

np. random. seed(2026)
torch. manual_seed(2026)
quarterly = np. array([
 5031,4697,5959,5665, 5138,5454,5618,5529,
 4991,4995,5467,5541, 4868,5413,5308,5631,
 4759,4939,5378,5270
], dtype=float)
weights = np. array([0.31, 0.32, 0.37])
monthly = []
for q in quarterly:
 v = np. rint(q * weights). astype(int)
 v[-1] += int(q - v. sum())
 monthly. extend(v. tolist())
monthly = np. asarray(monthly, dtype=np. float32)

mean, std = monthly[:48]. mean(), monthly[:48]. std()
z = (monthly - mean) / std
seq_len = 6
X, Y = [], []
for i in range(seq_len, len(z)):
 X. append(z[i-seq_len: i])
 Y. append(z[i])
X = torch. tensor(np. array(X)[:,:, None], dtype=torch. float32)
Y = torch. tensor(np. array(Y)[:, None], dtype=torch. float32)

split = 42
Xtr, Xte = X[: split], X[split:]
Ytr, Yte = Y[: split], Y[split:]
```

```
class LSTMForecast(nn. Module):
 def __init__(self, hidden=32):
 super().__init__()
 self. lstm = nn. LSTM(input_size=1, hidden_size=hidden, batch_first=True)
 self. fc = nn. Linear(hidden, 1)
 def forward(self, x):
 out, _ = self. lstm(x)
 return self. fc(out[:, -1,:])

model = LSTMForecast(32)
opt = torch. optim. Adam(model. parameters(), lr=0.01)
loss_fn = nn. MSELoss()
history = []
for epoch in range(350):
 model. train(); opt. zero_grad()
 pred = model(Xtr)
 loss = loss_fn(pred, Ytr)
 loss. backward()
 torch. nn. utils. clip_grad_norm_(model. parameters(), 1.0)
 opt. step()
 history. append(loss. item())

model. eval()
with torch. no_grad():
 pred_te = model(Xte). squeeze(). numpy() * std + mean
 real_te = Yte. squeeze(). numpy() * std + mean
mae = np. mean(np. abs(real_te - pred_te))
print("MAE mensual derivado =", round(float(mae), 2))

plt. figure(figsize=(8,4))
plt. plot(real_te, marker="o", label="Real derivado")
plt. plot(pred_te, marker="o", label="LSTM")
plt. legend(); plt. title("Evaluación de LSTM"); plt. tight_layout(); plt. show()
```

Preguntas para el laboratorio

• ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?

• ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?

• ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?

• ¿Qué información de campo agregaría para reducir incertidumbre?

• ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO VI

Auditorías de Seguridad Vial y Modelos de Evaluación del Riesgo

AUTOR: Cañarte Baque George Alfredo

EDITORIAL RUNAIKI

6 Auditorías de Seguridad Vial y Modelos de Evaluación del Riesgo

La auditoría vial identifica deficiencias antes de que se traduzcan en daño humano. Cuando sus observaciones se estructuran como variables, la inspección puede complementarse con modelos de probabilidad, árboles y ensamblados. La finalidad es priorizar medidas que reduzcan energía de impacto, exposición y posibilidad de conflicto. La auditoría cuantificada gana valor cuando los hallazgos físicos: velocidad, fricción, geometría, señalización y protección lateral: se vinculan con probabilidad, severidad y exposición sin sustituir la inspección profesional (PIARC, 2025; International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

Resultados de aprendizaje

- Reconocer el alcance preventivo de una auditoría de seguridad vial y sus requisitos de independencia técnica.
- Estructurar indicadores de riesgo a partir de geometría, operación, condición superficial y usuarios vulnerables.
- Aplicar logit, probit, árboles y Random Forest para apoyar la identificación de tramos críticos.
- Integrar evidencia de campo y resultados predictivos sin sustituir la revisión profesional.
- Priorizar intervenciones mediante análisis multicriterio con sensibilidad a pesos y presupuesto.

6.1 Concepto y alcance de auditorías viales en ingeniería civil

Desde la ingeniería civil, el estudio de concepto y alcance de auditorías viales en ingeniería civil exige algo más que ejecutar una rutina estadística. El análisis combina etapa del proyecto, usuarios, riesgos y evidencia de campo y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Breiman, 2001; Baque Parrales et al., 2022).

En proyectos nuevos y vías en operación, la auditoría independiente debe sustentarse en criterios explícitos, registros versionados y reglas de decisión

previamente acordadas. La prioridad consiste en detectar oportunidades de reducción del daño antes de que se materialice. El desempeño estadístico se valora únicamente cuando mejora esa decisión. El resultado adquiere valor práctico cuando puede contrastarse con evidencia independiente, ya provenga de inspección, normativa, medición de campo o un periodo posterior (Moreno-Ponce et al., 2024; International Road Assessment Programme, 2024).

6.1.1 Propósito preventivo de una auditoría de seguridad vial

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de etapa del proyecto, usuarios, riesgos y evidencia de campo se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Baque Parrales et al., 2022; Federal Highway Administration, 2006).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la auditoría independiente de seguridad vial se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La importancia asignada por un árbol o un bosque no demuestra causalidad; debe contrastarse con geometría, condición superficial y evidencia procedente de la auditoría. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Moreno-Ponce et al., 2024; PIARC, 2025).

La utilidad del análisis se hace evidente cuando permite detectar oportunidades de reducción del daño antes de que se materialice. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En proyectos nuevos y vías en operación, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2024; Breiman, 2001).

6.1.2 Etapas del proyecto y oportunidad de la revisión

La evaluación se sustenta en etapa del proyecto, usuarios, riesgos y evidencia de campo, seleccionados por su relación directa con la decisión de ingeniería.

El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (International Road Assessment Programme, 2024; Baque Parrales et al., 2022).

La estimación utiliza auditoría independiente de seguridad vial, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. La importancia asignada por un árbol o un bosque no demuestra causalidad; debe contrastarse con geometría, condición superficial y evidencia procedente de la auditoría. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (PIARC, 2025; International Road Assessment Programme, 2024).

La aplicación del procedimiento cobra sentido al apoyar la decisión de detectar oportunidades de reducción del daño antes de que se materialice. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En proyectos nuevos y vías en operación, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (PIARC, 2025; Breiman, 2001).

Para visualizar matriz de prioridad para hallazgos de auditoría vial, la Figura 6.1 presenta una representación técnica que complementa la explicación y facilita su reproducción.

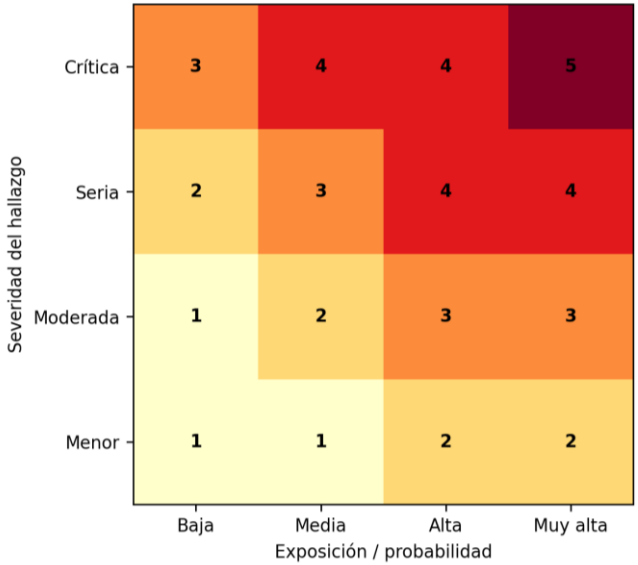


Figura 6.1. Matriz de prioridad para hallazgos de auditoría vial.

Nota. Elaboración propia. La escala combina exposición/probabilidad y severidad para organizar la revisión técnica.

6.1.3 Independencia, competencias y composición del equipo auditor

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean etapa del proyecto, usuarios, riesgos y evidencia de campo y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Breiman, 2001; Baque Parrales et al., 2022).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la auditoría independiente de seguridad vial se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La importancia asignada por un árbol o un bosque no demuestra causalidad; debe contrastarse con geometría, condición superficial y evidencia procedente de la auditoría. La evaluación se centra en desempeño,

estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (International Road Assessment Programme, 2024; Federal Highway Administration, 2006).

La aplicación del procedimiento cobra sentido al apoyar la decisión de detectar oportunidades de reducción del daño antes de que se materialice. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En proyectos nuevos y vías en operación, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (International Road Assessment Programme, 2024; Breiman, 2001).

6.1.4 Hallazgo, evidencia y recomendación técnicamente verificable

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean etapa del proyecto, usuarios, riesgos y evidencia de campo y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Baque Parrales et al., 2022; PIARC, 2025).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora la auditoría independiente de seguridad vial. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. La importancia asignada por un árbol o un bosque no demuestra causalidad; debe contrastarse con geometría, condición superficial y evidencia procedente de la auditoría. El criterio de selección se define antes de observar el conjunto reservado para prueba (Moreno-Ponce et al., 2024; International Road Assessment Programme, 2024).

La aplicación del procedimiento cobra sentido al apoyar la decisión de detectar oportunidades de reducción del daño antes de que se materialice. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En proyectos nuevos y vías en operación, se recomienda informar el valor central, un rango de

incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Moreno-Ponce et al., 2024; Baque Parrales et al., 2022).

Ejercicio 6.1. Auditorías de seguridad vial en ingeniería civil

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con geometría, señalización, visibilidad, drenaje, velocidad y usuarios vulnerables. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Aplique una lista de verificación a 12 segmentos, documente evidencia fotográfica y asigne gravedad, exposición y tratabilidad.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: programa de inspección priorizada con responsable y plazo. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

6.2 Metodologías internacionales y estructuración de indicadores

Para trabajar con metodologías internacionales y estructuración de indicadores de forma reproducible, se establece primero la unidad de análisis y la pregunta que orienta el procedimiento. Se consideran atributos de vía, severidad, exposición e indicadores normalizados, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas,

la discrepancia se documenta antes de realizar cualquier ajuste (Baque Parrales et al., 2022; Breiman, 2001).

En evaluaciones viales de Ecuador y Manabí, el Sistema Seguro, iRAP y el HSM deben articularse mediante criterios explícitos, información versionada y reglas de decisión definidas previamente. El propósito es usar marcos compatibles sin mezclar escalas de forma incorrecta; una métrica aislada no constituye por sí sola un criterio de decisión. La utilidad profesional depende de que el resultado pueda verificarse mediante inspección, norma, medición independiente o datos de un periodo posterior no utilizado en el ajuste (Federal Highway Administration, 2006; PIARC, 2025).

6.2.1 Sistema Seguro como marco de lectura de hallazgos

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de atributos de vía, severidad, exposición e indicadores normalizados se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (PIARC, 2025; Federal Highway Administration, 2006).

La evaluación utiliza Sistema Seguro, iRAP y HSM, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Federal Highway Administration, 2006; International Road Assessment Programme, 2024).

Los resultados deben interpretarse con una finalidad explícita: usar marcos compatibles sin mezclar escalas de forma incorrecta. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En evaluaciones viales en Ecuador y Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una

predicción como si fuera evidencia de campo (Federal Highway Administration, 2006; Breiman, 2001).

6.2.2 iRAP y evaluación estandarizada de atributos de la vía

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean atributos de vía, severidad, exposición e indicadores normalizados y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (PIARC, 2025; Federal Highway Administration, 2006).

La estimación utiliza Sistema Seguro, iRAP y HSM, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Federal Highway Administration, 2006; International Road Assessment Programme, 2024).

El resultado es útil cuando permite usar marcos compatibles sin mezclar escalas de forma incorrecta. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En evaluaciones viales en Ecuador y Manabí, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Federal Highway Administration, 2006; Breiman, 2001).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible

para priorizar recursos de inspección (International Road Assessment Programme, 2024; Baque Parrales et al., 2022).

6.2.3 Highway Safety Manual y funciones de desempeño

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean atributos de vía, severidad, exposición e indicadores normalizados y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (PIARC, 2025; Breiman, 2001).

La estimación utiliza Sistema Seguro, iRAP y HSM, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. La importancia asignada por un árbol o un bosque no demuestra causalidad; debe contrastarse con geometría, condición superficial y evidencia procedente de la auditoría. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Federal Highway Administration, 2006; Moreno-Ponce et al., 2024).

La aplicación del procedimiento cobra sentido al apoyar la decisión de usar marcos compatibles sin mezclar escalas de forma incorrecta. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En evaluaciones viales en Ecuador y Manabí, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Federal Highway Administration, 2006; PIARC, 2025).

Para ordenar marcos utilizados en evaluación de seguridad vial, la Tabla 6.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 6-1. Marcos utilizados en evaluación de seguridad vial

Marco	Unidad de lectura	Aporte
Sistema Seguro	Sistema vial	Tolerancia al error
iRAP	Atributos de carretera	Clasificación estandarizada
HSM	Segmento/intersección	Frecuencia esperada
RSA	Proyecto/corredor	Hallazgos preventivos

Nota. Elaboración propia con base en PIARC (2025), iRAP (2024), FHWA (2006) y AASHTO (2010).

6.2.4 Armonización de indicadores para evitar doble conteo

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de atributos de vía, severidad, exposición e indicadores normalizados se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Breiman, 2001; Baque Parrales et al., 2022).

Para mantener la reproducibilidad, el análisis incorpora Sistema Seguro, iRAP y HSM dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. El criterio de selección se define antes de observar el conjunto reservado para prueba (International Road Assessment Programme, 2024; PIARC, 2025).

En términos de ingeniería, el análisis orienta la decisión de usar marcos compatibles sin mezclar escalas de forma incorrecta. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En evaluaciones viales en Ecuador y Manabí, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida

física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (International Road Assessment Programme, 2024; Breiman, 2001).

Una comprobación adicional consiste en retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Moreno-Ponce et al., 2024; Baque Parrales et al., 2022).

Ejercicio 6.2. Metodologías internacionales e indicadores de auditoría

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con indicadores de auditoría, severidad y frecuencia. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Homologue escalas de distintas guías internacionales y construya un índice de 0 a 100 sin mezclar variables duplicadas.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: ficha de indicadores con definición, unidad, fuente y regla de puntuación. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

6.3 Modelación del riesgo vial mediante regresión logística: logit y probit

Para analizar modelación del riesgo vial mediante regresión logística: logit y probit con rigor, primero se estructura la evidencia y luego se selecciona el procedimiento de modelación más apropiado. La base reúne respuesta binaria, predictores y probabilidad estimada y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Federal Highway Administration, 2006; Baque Parrales et al., 2022).

El procedimiento adoptado emplea logit y probit y vincula cada salida con una acción verificable: clasificar riesgo con calibración y umbrales explícitos. La aplicación en segmentos con presencia o ausencia de condición crítica obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (PIARC, 2025; International Road Assessment Programme, 2024).

6.3.1 Probabilidad de riesgo y función logística

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean respuesta binaria, predictores y probabilidad estimada y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Baque Parrales et al., 2022; Breiman, 2001).

$$P(Y = 1 \mid \mathbf{x}) = \frac{1}{1 + e^{-(\beta_0 + \boldsymbol{\beta}^T \mathbf{x})}}$$

La función logística transforma el predictor lineal en una probabilidad entre cero y uno. Cada coeficiente expresa un cambio en el logaritmo de las odds, por lo que su interpretación debe acompañarse de unidades y rangos plausibles de las variables.

Para mantener la reproducibilidad, el análisis incorpora logit y probit dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. El criterio de selección se define antes de observar el conjunto reservado para prueba (Moreno-Ponce et al., 2024; Baque Parrales et al., 2022).

6.3.2 Odds, razones de momios e interpretación de coeficientes

Para abordar el problema se integran respuesta binaria, predictores y probabilidad estimada, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (International Road Assessment Programme, 2024; Federal Highway Administration, 2006).

La estimación utiliza logit y probit, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Los indicadores de importancia del modelo describen capacidad predictiva, no causalidad. La explicación técnica se completa con la inspección geométrica, superficial y funcional. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (PIARC, 2025; International Road Assessment Programme, 2024).

Para la ingeniería, el análisis resulta pertinente cuando permite clasificar riesgo con calibración y umbrales explícitos. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En segmentos con presencia o ausencia de condición crítica, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (PIARC, 2025; Breiman, 2001).

6.3.3 Modelo probit como alternativa de enlace

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de respuesta binaria, predictores y probabilidad estimada se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (PIARC, 2025; Baque Parrales et al., 2022).

$$P(Y = 1 \mid \mathbf{x}) = \Phi(\beta_0 + \boldsymbol{\beta}^T \mathbf{x})$$

En el modelo probit, Φ representa la función de distribución acumulada normal estándar. Logit y probit suelen producir clasificaciones semejantes, aunque difieren en la forma del enlace y en la escala de sus coeficientes.

Para sostener la interpretación, el procedimiento basado en logit y probit se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. Los indicadores de importancia del modelo describen capacidad predictiva, no causalidad. La explicación técnica se completa con la inspección geométrica, superficial y funcional. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Federal Highway Administration, 2006; PIARC, 2025).

6.3.4 Discriminación, calibración y elección del umbral de decisión

El desarrollo del apartado requiere trabajar con respuesta binaria, predictores y probabilidad estimada. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (International Road Assessment Programme, 2024; Breiman, 2001).

La estimación utiliza logit y probit, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Una alta

importancia predictiva no implica un efecto causal. Su interpretación requiere revisar la geometría, el estado superficial y los hallazgos de campo. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (PIARC, 2025; Moreno-Ponce et al., 2024).

Ejercicio guiado 6.3. Regresión logística para evaluación del riesgo

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con evento grave 0-1, V85, radio, iluminación, lluvia y usuario vulnerable. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Divida datos, ajuste regresión logística, convierta coeficientes a odds ratios y evalúe AUC, sensibilidad y especificidad.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: probabilidad de evento grave por segmento y explicación de los factores modificables. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

6.4 Evaluación funcional y geométrica de carreteras con variables predictoras

La comprensión de evaluación funcional y geométrica de carreteras con variables predictoras requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran radio, pendiente, IRI, fricción, señalización, iluminación y V85; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar

tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Moreno-Ponce et al., 2024; International Road Assessment Programme, 2024).

En la práctica, evaluación funcional y geométrica permite organizar la información para relacionar deficiencias observables con riesgo operacional. El análisis se plantea para Jipijapa–La Pila, El Carmen y corredores urbanos, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Baque Parrales et al., 2022; Federal Highway Administration, 2006).

6.4.1 Geometría horizontal y vertical en la evaluación funcional

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran radio, pendiente, IRI, fricción, señalización, iluminación y V85, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Moreno-Ponce et al., 2024; Federal Highway Administration, 2006).

Para mantener la reproducibilidad, el análisis incorpora evaluación funcional y geométrica dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. El criterio de selección se define antes de observar el conjunto reservado para prueba (Breiman, 2001; PIARC, 2025).

Los resultados deben interpretarse con una finalidad explícita: relacionar deficiencias observables con riesgo operacional. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a

medirse después de una intervención. En Jipijapa–La Pila, El Carmen y corredores urbanos, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Breiman, 2001; Moreno-Ponce et al., 2024).

6.4.2 IRI, macrotextura y fricción como evidencias de condición

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de radio, pendiente, IRI, fricción, señalización, iluminación y V85 se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Breiman, 2001; Federal Highway Administration, 2006).

Para mantener la reproducibilidad, el análisis incorpora evaluación funcional y geométrica dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. El criterio de selección se define antes de observar el conjunto reservado para prueba (International Road Assessment Programme, 2024; Breiman, 2001).

El resultado es útil cuando permite relacionar deficiencias observables con riesgo operacional. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En Jipijapa–La Pila, El Carmen y corredores urbanos, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

6.4.3 Señalización, visibilidad e iluminación del corredor

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean radio, pendiente, IRI, fricción, señalización, iluminación y V85 y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (PIARC, 2025; International Road Assessment Programme, 2024).

La estimación utiliza evaluación funcional y geométrica, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Federal Highway Administration, 2006; Breiman, 2001).

Los resultados deben interpretarse con una finalidad explícita: relacionar deficiencias observables con riesgo operacional. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En Jipijapa–La Pila, El Carmen y corredores urbanos, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Federal Highway Administration, 2006; Baque Parrales et al., 2022).

6.4.4 Velocidad V85 y coherencia entre diseño y operación

En este apartado, la evidencia se organiza a partir de radio, pendiente, IRI, fricción, señalización, iluminación y V85. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de

corrección aplicada a valores imposibles o incompletos (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

La estimación utiliza evaluación funcional y geométrica, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. La contribución de una variable al modelo debe leerse como evidencia predictiva y no como causa demostrada; la auditoría de campo sigue siendo necesaria. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (PIARC, 2025; Baque Parrales et al., 2022).

Ejercicio 6.4. Evaluación funcional y geométrica de carreteras

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con radio, peralte, pendiente, distancia de visibilidad, ancho y fricción. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule criterios geométricos y compárelos con valores de operación; marque incumplimientos y vincúlelos con siniestros o conflictos observados.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: ficha técnica de geometría con prioridad de corrección. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

6.5 Identificación de tramos críticos mediante árboles de decisión y Random Forest

La comprensión de identificación de tramos críticos mediante árboles de decisión y random forest requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran variables de inspección, clases y relevancia predictiva; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Moreno-Ponce et al., 2024; Federal Highway Administration, 2006).

En bases con relaciones no lineales e interacciones, los árboles de decisión y Random Forest requieren parámetros documentados, datos versionados y una regla de decisión definida antes de evaluar resultados. El análisis se orienta a localizar tramos con combinaciones críticas de atributos, mientras que la métrica funciona como criterio de apoyo y no como un fin en sí misma. La aplicación técnica requiere una comprobación externa mediante inspección, norma, medición independiente o información posterior al periodo de ajuste (Baque Parrales et al., 2022; Breiman, 2001).

6.5.1 Particiones recursivas y reglas interpretables de decisión

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de variables de inspección, clases y relevancia predictiva se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Breiman, 2001; Moreno-Ponce et al., 2024).

$$G = 1 - \sum_{k=1}^{K} p_k^2$$

El índice de Gini G mide la impureza de un nodo. Una partición es favorable cuando reduce esa impureza y separa observaciones con respuestas distintas de manera suficientemente estable.

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando los árboles de decisión y Random Forest como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. Una alta importancia predictiva no implica un efecto causal. Su interpretación requiere revisar la geometría, el estado superficial y los hallazgos de campo. El criterio de selección se define antes de observar el conjunto reservado para prueba (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de localizar tramos con combinaciones críticas de atributos. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En bases con relaciones no lineales e interacciones, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (International Road Assessment Programme, 2024; Federal Highway Administration, 2006).

6.5.2 Bosques aleatorios y reducción de varianza

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de variables de inspección, clases y relevancia predictiva se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Baque Parrales et al., 2022; International Road Assessment Programme, 2024).

La interpretación requiere una comprobación externa; la salida obtenida mediante los árboles de decisión y Random Forest se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. Una alta importancia predictiva no implica un efecto causal. Su interpretación requiere revisar la geometría, el estado superficial y los hallazgos de campo. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Moreno-Ponce et al., 2024; Baque Parrales et al., 2022).

Los resultados aportan valor técnico cuando permiten localizar tramos con combinaciones críticas de atributos. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En bases con relaciones no lineales e interacciones, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2024; Federal Highway Administration, 2006).

Para visualizar importancia relativa ilustrativa de variables en un random forest vial, la Figura 6.2 presenta una representación técnica que complementa la explicación y facilita su reproducción.

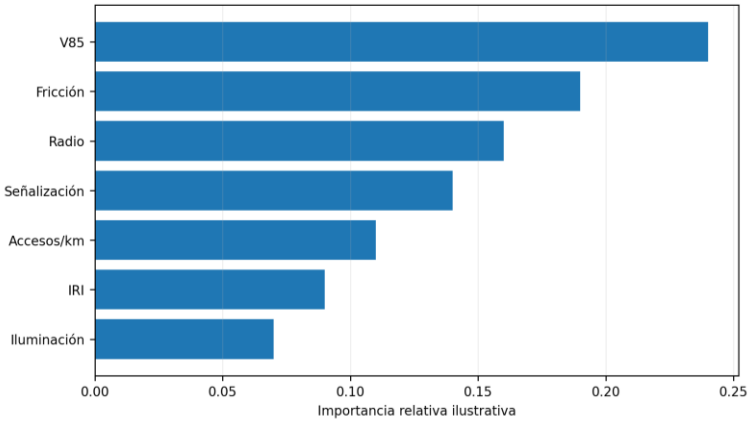


Figura 6.2. Importancia relativa ilustrativa de variables en un Random Forest vial.

Nota. Elaboración propia. Los valores se utilizan con fines didácticos y no corresponden a una carretera específica.

6.5.3 Importancia de variables y dependencia parcial

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean variables de inspección, clases y relevancia predictiva y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios

de codificación, precisión o cobertura entre campañas de levantamiento (Breiman, 2001; International Road Assessment Programme, 2024).

La interpretación requiere una comprobación externa; la salida obtenida mediante los árboles de decisión y Random Forest se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. Una alta importancia predictiva no implica un efecto causal. Su interpretación requiere revisar la geometría, el estado superficial y los hallazgos de campo. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

Los resultados aportan valor técnico cuando permiten localizar tramos con combinaciones críticas de atributos. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En bases con relaciones no lineales e interacciones, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (International Road Assessment Programme, 2024; Baque Parrales et al., 2022).

6.5.4 Datos desbalanceados y validación de tramos críticos

El desarrollo del apartado requiere trabajar con variables de inspección, clases y relevancia predictiva. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

La estimación utiliza árboles de decisión y Random Forest, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Una alta importancia predictiva no implica un efecto causal. Su interpretación requiere revisar la geometría, el estado superficial y los hallazgos de campo. Para repetir el análisis deben registrarse la versión de

la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (PIARC, 2025; Federal Highway Administration, 2006).

En auditoría vial, una clasificación binaria es menos informativa que la probabilidad estimada de riesgo alto por corredor. La Figura 6.3 compara dos salidas predictivas obtenidas con Random Forest y regresión logística sobre datos simulados agrupados por corredor.

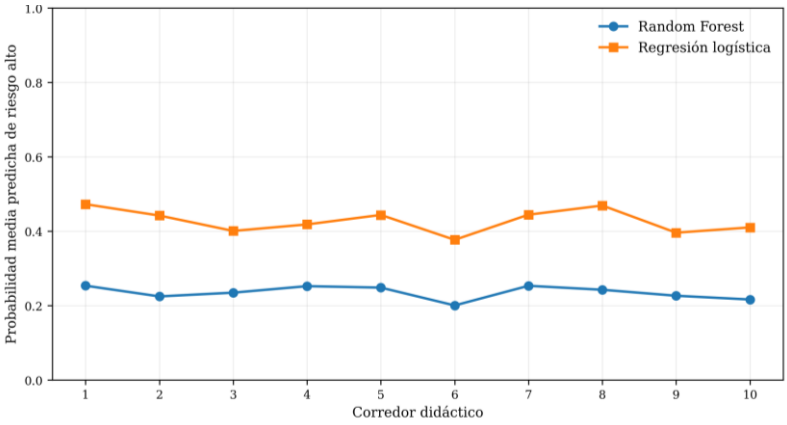


Figura 6.3. Probabilidad media predicha de riesgo alto por corredor.

Nota. Elaboración propia a partir del Laboratorio Python 6; datos simulados y validación agrupada por corredor.

La coincidencia entre modelos puede reforzar una prioridad preliminar, mientras que diferencias amplias requieren revisar variables, balance de clases y evidencia de auditoría. La probabilidad constituye un filtro para inspección; la decisión de intervención conserva la comprobación técnica del tramo.

Ejercicio 6.5. Árboles de decisión y Random Forest para tramos críticos

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con condición del pavimento, V85, radio, drenaje, señalización, TPDA y riesgo. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Entrene árbol y Random Forest, reserve datos de prueba, calcule importancia por permutación y revise errores de los segmentos críticos.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: ranking de variables y listado de tramos de mayor probabilidad. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

6.6 Integración de modelos predictivos en auditorías viales

El estudio de integración de modelos predictivos en auditorías viales exige ordenar primero la evidencia y seleccionar después el algoritmo. La base reúne hallazgos, probabilidades, evidencia fotográfica y medidas y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (International Road Assessment Programme, 2024; Breiman, 2001).

Desde una perspectiva operativa, la integración entre auditoría y modelo ayuda a estructurar la información y a ordenar inspecciones sin automatizar la decisión final. El análisis se plantea para programas de seguridad vial municipales y provinciales, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Breiman, 2001; Baque Parrales et al., 2022).

6.6.1 Vinculación entre hallazgo de campo y estimación probabilística

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de hallazgos, probabilidades, evidencia fotográfica y medidas se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Baque Parrales et al., 2022; Breiman, 2001).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la integración entre auditoría y modelo se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La importancia asignada por un árbol o un bosque no demuestra causalidad; debe contrastarse con geometría, condición superficial y evidencia procedente de la auditoría. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Moreno-Ponce et al., 2024; Baque Parrales et al., 2022).

El resultado cuantitativo se convierte en una decisión operativa cuando permite ordenar inspecciones sin automatizar la decisión final. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En programas de seguridad vial municipales y provinciales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Moreno-Ponce et al., 2024; Federal Highway Administration, 2006).

6.6.2 Ficha de auditoría enriquecida con evidencia predictiva

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean hallazgos, probabilidades, evidencia fotográfica y medidas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Breiman, 2001; Federal Highway Administration, 2006).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante la integración entre auditoría y modelo se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La importancia asignada por un árbol o un bosque no demuestra causalidad; debe contrastarse con geometría, condición superficial y evidencia procedente de la auditoría. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (International Road Assessment Programme, 2024; Breiman, 2001).

La salida del análisis adquiere sentido cuando se vincula con una finalidad concreta: ordenar inspecciones sin automatizar la decisión final. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En programas de seguridad vial municipales y provinciales, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Moreno-Ponce et al., 2024; Federal Highway Administration, 2006).

6.6.3 Priorización de inspecciones sin sustituir el juicio profesional

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran hallazgos, probabilidades, evidencia fotográfica y medidas, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Federal Highway Administration, 2006; PIARC, 2025).

La estimación utiliza integración auditoría-modelo, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. Una alta importancia predictiva no implica un efecto causal. Su interpretación requiere revisar la geometría, el estado superficial y los hallazgos de campo. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Baque Parrales et al., 2022; PIARC, 2025).

Los resultados aportan valor técnico cuando permiten ordenar inspecciones sin automatizar la decisión final. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En programas de seguridad vial municipales y provinciales, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Baque Parrales et al., 2022; Breiman, 2001).

La sensibilidad del resultado se estudia mediante variaciones controladas de los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (PIARC, 2025; Baque Parrales et al., 2022).

6.6.4 Verificación posterior de la medida aplicada

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de hallazgos, probabilidades, evidencia fotográfica y medidas se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Breiman, 2001; PIARC, 2025).

La aplicación metodológica de la integración entre auditoría y modelo sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. Los indicadores de importancia del modelo describen capacidad predictiva, no causalidad. La explicación técnica se completa con la inspección geométrica, superficial y funcional. El criterio de selección se define antes de observar el conjunto reservado para prueba (International Road Assessment Programme, 2024; Breiman, 2001).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: ordenar inspecciones sin automatizar la decisión final. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En programas de seguridad vial municipales y provinciales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

Como verificación complementaria, se propone retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Moreno-Ponce et al., 2024; Federal Highway Administration, 2006).

Ejercicio 6.6. Integración de modelos predictivos en auditorías viales

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con resultado de auditoría, probabilidad del modelo y costo de inspección. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.

2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Combine evidencia física y probabilidad predictiva mediante una matriz de dos ejes; no permita que el algoritmo elimine una observación de seguridad grave.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: cartera de inspecciones ordenada por severidad y evidencia acumulada. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

6.7 Diseño de intervenciones optimizadas mediante análisis multicriterio

En diseño de intervenciones optimizadas mediante análisis multicriterio, la elección del algoritmo ocupa un segundo plano frente a la organización y calidad de la evidencia disponible. La base reúne riesgo, costo, población expuesta, factibilidad y beneficio y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Federal Highway Administration, 2006; PIARC, 2025).

La aplicación práctica de la normalización, la ponderación y TOPSIS permite disponer la información de manera que sea posible priorizar medidas bajo recursos limitados. El análisis se plantea para carteras de señalización, fricción, drenaje y protección peatonal, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (PIARC, 2025; Breiman, 2001).

6.7.1 Definición de criterios técnicos, sociales y económicos

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de riesgo, costo, población expuesta, factibilidad y beneficio se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Breiman, 2001; PIARC, 2025).

La aplicación metodológica de la normalización, la ponderación y TOPSIS sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. Los indicadores de importancia del modelo describen capacidad predictiva, no causalidad. La explicación técnica se completa con la inspección geométrica, superficial y funcional. El criterio de selección se define antes de observar el conjunto reservado para prueba (International Road Assessment Programme, 2024; Breiman, 2001).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: priorizar medidas bajo recursos limitados. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En carteras de señalización, fricción, drenaje y protección peatonal, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (International Road Assessment Programme, 2024; Moreno-Ponce et al., 2024).

6.7.2 Asignación de pesos y normalización de escalas

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de riesgo, costo, población expuesta, factibilidad y beneficio se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Baque Parrales et al., 2022; Federal Highway Administration, 2006).

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante la normalización, la ponderación y TOPSIS frente a una referencia ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. Los indicadores de importancia del modelo describen capacidad predictiva, no causalidad. La explicación técnica se completa con la inspección geométrica, superficial y funcional. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Moreno-Ponce et al., 2024; PIARC, 2025).

Para la ingeniería, el análisis resulta pertinente cuando permite priorizar medidas bajo recursos limitados. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En carteras de señalización, fricción, drenaje y protección peatonal, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Moreno-Ponce et al., 2024; Breiman, 2001).

6.7.3 TOPSIS y distancia a soluciones de referencia

Para abordar el problema se integran riesgo, costo, población expuesta, factibilidad y beneficio, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Federal Highway Administration, 2006; International Road Assessment Programme, 2024).

$$C_i^* = \frac{S_i^-}{S_i^+ + S_i^-}$$

El coeficiente C_i* aumenta cuando una alternativa se aproxima a la solución ideal positiva y se aleja de la negativa. Los resultados dependen de la normalización y de los pesos, por lo que deben someterse a análisis de sensibilidad.

Para sostener la interpretación, el procedimiento basado en la normalización, la ponderación y TOPSIS se complementa con una comprobación directa

desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. Los indicadores de importancia del modelo describen capacidad predictiva, no causalidad. La explicación técnica se completa con la inspección geométrica, superficial y funcional. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Baque Parrales et al., 2022; Breiman, 2001).

Para la ingeniería, el análisis resulta pertinente cuando permite priorizar medidas bajo recursos limitados. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En carteras de señalización, fricción, drenaje y protección peatonal, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Baque Parrales et al., 2022; Federal Highway Administration, 2006).

Para ordenar matriz de criterios para priorizar intervenciones, la Tabla 6.2 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 6-2. Matriz de criterios para priorizar intervenciones

Criterio	Indicador	Sentido
Riesgo	Probabilidad × severidad	Maximizar reducción
Costo	USD estimado	Minimizar
Usuarios expuestos	personas/día	Maximizar cobertura
Factibilidad	escala 1-5	Maximizar
Tiempo de ejecución	días	Minimizar

Nota. Elaboración propia. Los pesos deben someterse a sensibilidad y aprobación técnica.

6.7.4 Sensibilidad de la priorización ante presupuesto y pesos

El desarrollo del apartado requiere trabajar con riesgo, costo, población expuesta, factibilidad y beneficio. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse,

actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Moreno-Ponce et al., 2024; Baque Parrales et al., 2022).

La interpretación requiere una comprobación externa; la salida obtenida mediante la normalización, la ponderación y TOPSIS se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. Una alta importancia predictiva no implica un efecto causal. Su interpretación requiere revisar la geometría, el estado superficial y los hallazgos de campo. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Breiman, 2001; Federal Highway Administration, 2006).

Ejercicio guiado 6.7. Diseño de intervenciones mediante análisis multicriterio

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con costo, reducción esperada de riesgo, población beneficiaria, factibilidad y tiempo de ejecución. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Normalice criterios, asigne pesos documentados y aplique TOPSIS; repita el cálculo modificando cada peso en ±10 %.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: ranking robusto de intervenciones y registro de sensibilidad. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Laboratorio Python 6

El siguiente laboratorio implementa auditoría predictiva con random forest. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 6. Auditoría vial con Random Forest y regresión logística
# Requisitos: numpy, pandas, scikit-learn, matplotlib
import numpy as np
import pandas as pd
import matplotlib. pyplot as plt
from sklearn. model_selection import GroupKFold, cross_val_predict
from sklearn. ensemble import RandomForestClassifier
from sklearn. linear_model import LogisticRegression
from sklearn. preprocessing import StandardScaler
from sklearn. pipeline import make_pipeline
from sklearn. metrics import roc_auc_score

rng = np. random. default_rng(44)
n = 1000
df = pd. DataFrame({
 "corredor": rng. integers(1, 11, n),
 "velocidad": rng. normal(52, 15, n). clip(20, 100),
 "radio": rng. lognormal(np. log(180),.55, n). clip(35, 1200),
 "pendiente": rng. normal(3.5, 2.8, n). clip(0, 14),
 "berma": rng. integers(0, 2, n),
 "senalizacion": rng. integers(0, 2, n),
 "drenaje_deficiente": rng. binomial(1,.28, n),
 "tpda": rng. lognormal(np. log(520),.7, n). clip(40, 6000),
})
z = (-5.1 +.043*df. velocidad + 22/df. radio +.04*df. pendiente
 -.40*df. berma -.35*df. senalizacion
 +.55*df. drenaje_deficiente +.00014*df. tpda)
p = 1/(1+np. exp(-z))
df["riesgo_alto"] = (rng. random(n) < p). astype(int)

features = ["velocidad","radio","pendiente","berma","senalizacion",
 "drenaje_deficiente","tpda"]
X, y = df[features], df["riesgo_alto"]
groups = df["corredor"]
cv = GroupKFold(n_splits=5)

rf = RandomForestClassifier(
 n_estimators=400, min_samples_leaf=6,
 class_weight="balanced", random_state=42
)
prob_rf = cross_val_predict(rf, X, y, groups=groups, cv=cv,
 method="predict_proba")[:,1]
logit = make_pipeline(StandardScaler(), LogisticRegression(max_iter=3000,
```

```
 class_weight="balanced"))
prob_logit = cross_val_predict(logit, X, y, groups=groups, cv=cv,
 method="predict_proba")[:,1]
print("AUC RF:", round(roc_auc_score(y, prob_rf),3))
print("AUC Logit:", round(roc_auc_score(y, prob_logit),3))

rf. fit(X, y)
imp = pd. Series(rf. feature_importances_, index=features). sort_values()
imp. plot. barh(title="Importancia de variables")
plt. tight_layout(); plt. show()
```

Preguntas para el laboratorio

• ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?

• ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?

• ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?

• ¿Qué información de campo agregaría para reducir incertidumbre?

• ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

6.8 Referencias del capítulo 6

Federal Highway Administration. (2006). Road safety audit guidelines (FHWA-SA-06-06). U. S. Department of Transportation.

PIARC. (2025). Road safety manual: A manual for practitioners and decision makers on implementing safe system infrastructure. World Road Association. https://roadsafety. piarc. org/

International Road Assessment Programme. (2024). Star rating for designs: Supporting safer road designs. iRAP.

American Association of State Highway and Transportation Officials. (2010). Highway safety manual (1st ed.). AASHTO.

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi. org/10.1023/A:1010933404324

Gutiérrez Lino, K. A. (2024). Análisis de la señalización horizontal y vertical de la vía Jipijapa-Sancán [Tesis de grado, Universidad Estatal del Sur de Manabí]. Repositorio Digital UNESUM. http://repositorio. unesum. edu. ec/handle/53000/6528

Baque Parrales, E. M., Moll Martínez, R., Moreno Ponce, L. A., & Solórzano Villegas, L. E. (2022). Evaluación de la accidentalidad en la zona urbana de Jipijapa por deficiencias en la señalización vertical. Pro Sciences: Revista de Producción, Ciencias e Investigación, 6(45), 74–83. https://doi. org/10.29018/issn.2588-1000vol6iss45.2022pp74-83

Moreno-Ponce, L. A., Álvarez-Álvarez, M. J., Delgado-Alvia, R. P., & Zavala-Vásquez, C. J. (2024). Importancia del coeficiente de fricción para el análisis de seguridad vial, en el tramo de carretera Jipijapa - La Pila de la provincia de Manabí. Revista Científica INGENIAR: Ingeniería, Tecnología e Investigación, 7(13, ed. esp.). https://doi. org/10.46296/ig. v7i13edespab.0188

Moreno-Ponce, L. A., Solórzano-Villegas, L. E., Álvarez-Álvarez, M. J., & Castro-Chilán, D. M. (2025). Road safety and user vulnerability in urban corridors: Empirical evidence from the city of Jipijapa. Revista Científica INGENIAR: Ingeniería, Tecnología e Investigación, 8(16), 394–417.

Moreno-Ponce, L. A., Solórzano-Villegas, L. E., Ponce-Reyes, F. S., & Quimiz-Tumbaco, A. V. (2024). Análisis del Índice de Regularidad Internacional para la Seguridad Vial en el Paso Lateral El Carmen–Manabí. Revista Científica INGENIAR: Ingeniería, Tecnología e Investigación, 7(13 ed. esp.), 70–82. https://doi. org/10.46296/ig. v7i13edespab.0186

Suárez Toala, R. A., & Moreno Ponce, L. A. (2026). Evaluación y diseño de infraestructura peatonal segura en entornos escolares: Estudio aplicado en Ecuador. Reincisol, 6(10), 1670. https://doi. org/10.65835/reincisol. V6(10)1670

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO VII

Drenaje Vial y Modelación Hidrológica en la Seguridad Vial

AUTOR: Álvarez Álvarez Martha Johnna; Regalado Jalca Julio Johnny

EDITORIAL RUNAIKI

7 Drenaje Vial y Modelación Hidrológica en la Seguridad Vial

El agua modifica fricción, visibilidad, estabilidad de taludes, capacidad de soporte y continuidad operacional. El drenaje vial debe entenderse como parte de la seguridad, no solo como una obra complementaria. El capítulo une hidrología de diseño, hidráulica de cunetas y alcantarillas, SIG y evaluación de riesgo. La cuarta edición de HEC-22 actualiza los procedimientos de drenaje vial y el manejo de escorrentía sobre la plataforma, mientras que la revisión de Dehnad y Yazdi relaciona película de agua, macrotextura, neumático y velocidad con el hidroplaneo (Federal Highway Administration, 2024; Dehnad & Yazdi, 2024).

Resultados de aprendizaje

- Relacionar principios de hidráulica e hidrología con la seguridad y funcionalidad de la carretera.
- Dimensionar conceptualmente cunetas y alcantarillas mediante caudal, capacidad y criterios de mantenimiento.
- Explicar cómo lámina de agua, fricción e hidroplaneo modifican el riesgo operacional.
- Construir escenarios lluvia–escorrentía y eventos extremos para infraestructura vial.
- Integrar DEM, inventarios y modelos hidrológicos dentro de un SIG de priorización.

7.1 Fundamentos de hidráulica aplicada a infraestructura vial

El estudio de fundamentos de hidráulica aplicada a infraestructura vial exige ordenar primero la evidencia y seleccionar después el algoritmo. La base reúne caudal, área, velocidad, pendiente y rugosidad y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (Federal Highway Administration, 2024; Chow, 1959).

Desde una perspectiva operativa, continuidad, energía y Manning ayuda a estructurar la información y a comprobar que el agua se transporte sin erosionar ni invadir la calzada. El análisis se plantea para drenaje de carreteras rurales y

periurbanas, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Chow, 1959; World Meteorological Organization, 2008).

7.1.1 Continuidad y conservación de masa en drenaje vial

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean caudal, área, velocidad, pendiente y rugosidad y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora continuidad, energía y Manning. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. El caudal de diseño debe compararse con la capacidad hidráulica tanto en condición limpia como bajo escenarios razonables de obstrucción, sedimentación o aumento de rugosidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Dehnad & Yazdi, 2024; Chow, 1959).

7.1.2 Energía específica y pérdidas en conducciones y canales

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean caudal, área, velocidad, pendiente y rugosidad y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Chow, 1959; Chow et al., 1988).

La estimación utiliza continuidad, energía y Manning, con parámetros registrados y una línea base de comparación. Antes de estimar parámetros se

documenta la estrategia de calibración y verificación; si el método aprende de los datos, la partición de entrenamiento, validación y prueba se establece por anticipado. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. El registro técnico debe incluir versión de datos, parámetros, semilla aleatoria cuando proceda y fecha de ejecución, de manera que el cálculo pueda reconstruirse (Federal Highway Administration, 2024; Chow et al., 1988).

La lectura ingenieril del resultado debe traducirse en una acción concreta: comprobar que el agua se transporte sin erosionar ni invadir la calzada. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En drenaje de carreteras rurales y periurbanas, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Federal Highway Administration, 2024; Dehnad & Yazdi, 2024).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Chow et al., 1988; Federal Highway Administration, 2024).

7.1.3 Ecuación de Manning para flujo uniforme

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de caudal, área, velocidad, pendiente y rugosidad se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (World Meteorological Organization, 2008; Moreno-Ponce et al., 2024).

$$Q = \frac{1}{n} A R^{2/3} S^{1/2}$$

En la ecuación de Manning, Q es el caudal, n la rugosidad, A el área hidráulica, R el radio hidráulico y S la pendiente de energía. Su aplicación requiere comprobar que las condiciones de flujo sean compatibles con los supuestos del modelo (Chow, 1959).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora continuidad, energía y Manning. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. El caudal de diseño debe compararse con la capacidad hidráulica tanto en condición limpia como bajo escenarios razonables de obstrucción, sedimentación o aumento de rugosidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chow et al., 1988; Federal Highway Administration, 2024).

7.1.4 Régimen crítico y control hidráulico en obras de paso

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de caudal, área, velocidad, pendiente y rugosidad se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (World Meteorological Organization, 2008; Chow, 1959).

Para mantener la reproducibilidad, el análisis incorpora continuidad, energía y Manning dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chow et al., 1988; World Meteorological Organization, 2008).

La lectura ingenieril del resultado debe traducirse en una acción concreta: comprobar que el agua se transporte sin erosionar ni invadir la calzada. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué

información falta antes de diseñar una obra. En drenaje de carreteras rurales y periurbanas, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Chow et al., 1988; Dehnad & Yazdi, 2024).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Dehnad & Yazdi, 2024; Federal Highway Administration, 2024).

Ejercicio 7.1. Hidráulica aplicada a infraestructura vial

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con caudal, pendiente, rugosidad de Manning, ancho y tirante. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule área hidráulica, perímetro mojado, radio hidráulico, velocidad y capacidad; verifique límites de erosión o sedimentación.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: sección hidráulica que transporte el caudal de diseño con margen operativo. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

7.2 Tipologías de drenaje superficial y subterráneo

Desde la ingeniería civil, el estudio de tipologías de drenaje superficial y subterráneo exige algo más que ejecutar una rutina estadística. El análisis combina cunetas, subdrenes, sumideros, alcantarillas y descargas y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Chow, 1959; World Meteorological Organization, 2008).

En vías con lluvia estacional y mantenimiento limitado, la clasificación funcional del drenaje debe apoyarse en criterios explícitos, inventarios versionados y reglas previas de decisión. La prioridad consiste en seleccionar la solución según fuente y trayectoria del agua. El desempeño estadístico se valora únicamente cuando mejora esa decisión. El resultado adquiere valor práctico cuando puede contrastarse con evidencia independiente, ya provenga de inspección, normativa, medición de campo o un periodo posterior (Chow et al., 1988; Federal Highway Administration, 2024).

7.2.1 Drenaje superficial de plataforma y espaldones

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean cunetas, subdrenes, sumideros, alcantarillas y descargas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Chow, 1959; World Meteorological Organization, 2008).

La estimación utiliza clasificación funcional del drenaje, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. El caudal de diseño debe compararse con la capacidad hidráulica tanto en condición limpia como bajo escenarios razonables de obstrucción, sedimentación o aumento de rugosidad. La reproducibilidad requiere conservar la versión de los datos, los

parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Federal Highway Administration, 2024; Dehnad & Yazdi, 2024).

7.2.2 Drenaje subterráneo y control de humedad en pavimentos

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean cunetas, subdrenes, sumideros, alcantarillas y descargas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2024; Chow, 1959).

Para mantener la reproducibilidad, el análisis incorpora clasificación funcional del drenaje dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. El criterio de selección se define antes de observar el conjunto reservado para prueba (Dehnad & Yazdi, 2024; Chow et al., 1988).

En términos de ingeniería, el análisis orienta la decisión de seleccionar la solución según fuente y trayectoria del agua. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En vías con lluvia estacional y mantenimiento limitado, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Dehnad & Yazdi, 2024; Moreno-Ponce et al., 2024).

Para visualizar trayectoria funcional del agua desde la calzada hasta una descarga segura, la Figura 7.1 presenta una representación técnica que complementa la explicación y facilita su reproducción.

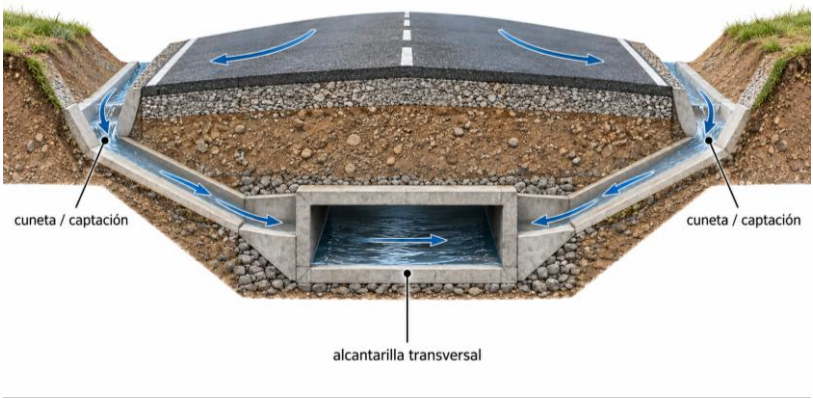


Figura 7.1. Trayectoria funcional del agua desde la calzada hasta una descarga segura.

Nota. Elaboración propia con base en principios de drenaje vial de FHWA (2024).

7.2.3 Alcantarillas y obras transversales de continuidad

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean cunetas, subdrenes, sumideros, alcantarillas y descargas y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2024; Chow et al., 1988).

Para mantener la reproducibilidad, el análisis incorpora clasificación funcional del drenaje dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. El criterio de selección se define antes de observar el conjunto reservado para prueba (Dehnad & Yazdi, 2024; Moreno-Ponce et al., 2024).

Los resultados deben interpretarse con una finalidad explícita: seleccionar la solución según fuente y trayectoria del agua. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En vías con lluvia estacional y mantenimiento limitado, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Dehnad & Yazdi, 2024; Federal Highway Administration, 2024).

7.2.4 Mantenimiento preventivo de cunetas, sumideros y descargas

En este apartado, la evidencia se organiza a partir de cunetas, subdrenes, sumideros, alcantarillas y descargas. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Federal Highway Administration, 2024; World Meteorological Organization, 2008).

La evaluación utiliza clasificación funcional del drenaje, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

Ejercicio 7.2. Drenaje superficial y subterráneo

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con tipo de obra, ubicación, diámetro o sección, estado, obstrucción y descarga. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Levante un inventario georreferenciado y clasifique fallas por capacidad, mantenimiento, erosión y continuidad de descarga.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: mapa de obras y agenda de limpieza o rehabilitación. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

7.3 Diseño hidráulico de cunetas y alcantarillas mediante modelos matemáticos

En diseño hidráulico de cunetas y alcantarillas mediante modelos matemáticos, la primera decisión consiste en precisar qué representa cada observación y qué pregunta deberá resolver el análisis. Se consideran intensidad de lluvia, cuenca, coeficiente de escorrentía y sección hidráulica, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Chow, 1959; Federal Highway Administration, 2024).

La sección articula método racional y ecuaciones de capacidad con una finalidad profesional concreta: dimensionar obras con margen y criterios de mantenimiento. En cunetas y alcantarillas de caminos secundarios, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método

imposible de mantener, explicar o verificar en terreno (Chow et al., 1988; World Meteorological Organization, 2008).

7.3.1 Método racional para estimar caudal de diseño

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de intensidad de lluvia, cuenca, coeficiente de escorrentía y sección hidráulica se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Chow, 1959; Moreno-Ponce et al., 2024).

$$Q = 0.278\, C\, i\, A$$

Con i en mm/h y A en km², el factor 0,278 permite obtener Q en m³/s. El coeficiente C representa la respuesta superficial de la cuenca y debe seleccionarse de acuerdo con cobertura, pendiente y condiciones de escorrentía.

El procedimiento combina método racional y ecuaciones de capacidad con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Federal Highway Administration, 2024; Chow, 1959).

La interpretación técnica orienta una acción verificable: dimensionar obras con margen y criterios de mantenimiento. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En cunetas y alcantarillas de caminos secundarios, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Federal Highway Administration, 2024; Chow et al., 1988).

7.3.2 Dimensionamiento de cunetas y control de velocidad erosiva

El desarrollo del apartado requiere trabajar con intensidad de lluvia, cuenca, coeficiente de escorrentía y sección hidráulica. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Federal Highway Administration, 2024; Chow et al., 1988).

El procedimiento combina método racional y ecuaciones de capacidad con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2024; Dehnad & Yazdi, 2024).

La interpretación se orienta a una decisión técnica definida: dimensionar obras con margen y criterios de mantenimiento. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En cunetas y alcantarillas de caminos secundarios, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

7.3.3 Funcionamiento de alcantarillas a entrada y salida

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran intensidad de lluvia, cuenca, coeficiente de escorrentía y sección hidráulica, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Federal Highway Administration, 2024; Chow et al., 1988).

La estimación utiliza método racional y ecuaciones de capacidad, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Moreno-Ponce et al., 2024; World Meteorological Organization, 2008).

La interpretación se orienta a una decisión técnica definida: dimensionar obras con margen y criterios de mantenimiento. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En cunetas y alcantarillas de caminos secundarios, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

Para ordenar variables para el dimensionamiento preliminar de drenaje, la Tabla 7.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 7-1. Variables para el dimensionamiento preliminar de drenaje

Variable	Símbolo	Unidad
Área aportante	A	ha o km²
Intensidad de lluvia	i	mm/h
Coeficiente de escorrentía	C	-
Caudal de diseño	Q	m³/s
Rugosidad Manning	n	s/m$^{1/3}$
Pendiente hidráulica	S	m/m

Nota. Elaboración propia con base en FHWA (2024) y Chow et al. (1988).

7.3.4 Resguardo, sedimentación y comprobación de capacidad

La evaluación se sustenta en intensidad de lluvia, cuenca, coeficiente de escorrentía y sección hidráulica, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Chow et al., 1988; World Meteorological Organization, 2008).

La estimación utiliza método racional y ecuaciones de capacidad, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. El caudal de diseño debe compararse con la capacidad hidráulica tanto en condición limpia como bajo escenarios razonables de obstrucción, sedimentación o aumento de rugosidad. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Chow, 1959; World Meteorological Organization, 2008).

Ejercicio 7.3. Diseño de cunetas y alcantarillas

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con área aportante, coeficiente C, intensidad, pendiente, rugosidad y geometría. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule caudal con el método racional y dimensione cuneta o alcantarilla con Manning; compruebe velocidad y resguardo.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: dimensionamiento completo con unidades y verificación hidráulica. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Ejemplo breve. Ejemplo de secuencia: primero Q=C·i·A; luego, con Q definido, resuelva la sección por Manning y compruebe velocidad, tirante y resguardo. No dimensione la cuneta antes de fijar el evento de diseño.

7.4 Relación entre escorrentía superficial y siniestralidad vial

La comprensión de relación entre escorrentía superficial y siniestralidad vial requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran profundidad de película, textura, fricción y velocidad; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Dehnad & Yazdi, 2024; World Meteorological Organization, 2008).

El procedimiento adoptado emplea mecanismos de hidroplaneo y frenado y vincula cada salida con una acción verificable: vincular la condición hidráulica con seguridad operacional. La aplicación en tramos húmedos y superficies con fricción reducida obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

7.4.1 Lámina de agua y reducción de contacto neumático-pavimento

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean profundidad de película, textura, fricción y velocidad y se revisa si cada medición

corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Chow, 1959; Federal Highway Administration, 2024).

La estimación utiliza mecanismos de hidroplaneo y frenado, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. La capacidad disponible se verifica frente al caudal de diseño considerando el funcionamiento normal y escenarios plausibles de reducción por sedimentos u obstrucciones. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Federal Highway Administration, 2024; Chow et al., 1988).

7.4.2 Hidroplaneo, velocidad y profundidad crítica

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de profundidad de película, textura, fricción y velocidad se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

La interpretación requiere una comprobación externa; la salida obtenida mediante los mecanismos de hidroplaneo y frenado se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Dehnad & Yazdi, 2024; Federal Highway Administration, 2024).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de vincular la condición hidráulica con seguridad operacional. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona

la red y permitir que otro profesional reconstruya el cálculo. En tramos húmedos y superficies con fricción reducida, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Dehnad & Yazdi, 2024; Chow et al., 1988).

7.4.3 Fricción disponible bajo lluvia y distancia de frenado

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean profundidad de película, textura, fricción y velocidad y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Chow, 1959; Chow et al., 1988).

$$d_b = \frac{v^2}{2gf}$$

Esta expresión simplificada muestra que la distancia de frenado d_b aumenta con el cuadrado de la velocidad v y disminuye cuando se reduce la fricción f. En diseño y auditoría deben incorporarse además tiempo de percepción-reacción, pendiente y condiciones reales del pavimento.

La estimación utiliza mecanismos de hidroplaneo y frenado, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Federal Highway Administration, 2024; World Meteorological Organization, 2008).

La interpretación se orienta a una decisión técnica definida: vincular la condición hidráulica con seguridad operacional. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En tramos húmedos y superficies con

fricción reducida, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Federal Highway Administration, 2024; Moreno-Ponce et al., 2024).

7.4.4 Vinculación entre fallas de drenaje y registros de siniestros

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de profundidad de película, textura, fricción y velocidad se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (World Meteorological Organization, 2008; Moreno-Ponce et al., 2024).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando los mecanismos de hidroplaneo y frenado como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Chow et al., 1988; Federal Highway Administration, 2024).

La interpretación técnica orienta una acción verificable: vincular la condición hidráulica con seguridad operacional. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En tramos húmedos y superficies con fricción reducida, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Chow et al., 1988; World Meteorological Organization, 2008).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible

para priorizar recursos de inspección (Dehnad & Yazdi, 2024; Moreno-Ponce et al., 2024).

Ejercicio 7.4. Escorrentía superficial y seguridad vial

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con lluvia, lámina de agua, fricción, V85, siniestros húmedos y secos. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Compare tasas de siniestros en condición húmeda y seca y estime una razón de riesgo; relacione el resultado con drenaje y fricción.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: identificación de segmentos donde el agua incrementa el riesgo operacional. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

7.5 Modelación hidrológica lluvia–escorrentía en carreteras

La comprensión de modelación hidrológica lluvia–escorrentía en carreteras requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran cuenca, IDF, tiempo de concentración e hidrograma; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Chow et al., 1988; Chow, 1959).

La sección articula modelación lluvia–escorrentía con una finalidad profesional concreta: estimar caudales compatibles con el área aportante. En cruces de drenaje en corredores viales, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (World Meteorological Organization, 2008; Dehnad & Yazdi, 2024).

7.5.1 Delimitación de cuencas aportantes a la carretera

El desarrollo del apartado requiere trabajar con cuenca, IDF, tiempo de concentración e hidrograma. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Federal Highway Administration, 2024; Chow et al., 1988).

El procedimiento combina modelación lluvia–escorrentía con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Moreno-Ponce et al., 2024; Dehnad & Yazdi, 2024).

La interpretación se orienta a una decisión técnica definida: estimar caudales compatibles con el área aportante. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En cruces de drenaje en corredores viales, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la

disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Chow, 1959; Chow et al., 1988).

7.5.2 Curvas IDF y selección del periodo de retorno

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran cuenca, IDF, tiempo de concentración e hidrograma, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chow et al., 1988; Dehnad & Yazdi, 2024).

Para sostener la interpretación, el procedimiento basado en modelación lluvia–escorrentía se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. La capacidad disponible se verifica frente al caudal de diseño considerando el funcionamiento normal y escenarios plausibles de reducción por sedimentos u obstrucciones. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Chow, 1959; Federal Highway Administration, 2024).

7.5.3 Tiempo de concentración y respuesta hidrológica

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean cuenca, IDF, tiempo de concentración e hidrograma y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2024; Chow, 1959).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando modelación lluvia–escorrentía como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta

el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Dehnad & Yazdi, 2024; World Meteorological Organization, 2008).

La interpretación técnica orienta una acción verificable: estimar caudales compatibles con el área aportante. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de inspección y los datos que todavía deben levantarse. En cruces de drenaje en corredores viales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Dehnad & Yazdi, 2024; Moreno-Ponce et al., 2024).

7.5.4 Hidrogramas de diseño para obras viales

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran cuenca, IDF, tiempo de concentración e hidrograma, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Dehnad & Yazdi, 2024; Chow, 1959).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando modelación lluvia–escorrentía como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. La comprobación hidráulica no se limita a la sección limpia; también incorpora pérdidas de capacidad por obstrucción, sedimentación y cambios de rugosidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (World Meteorological Organization, 2008; Dehnad & Yazdi, 2024).

La interpretación técnica orienta una acción verificable: estimar caudales compatibles con el área aportante. La utilidad del informe depende de vincular cada resultado con una condición física verificable, una responsabilidad de

inspección y los datos que todavía deben levantarse. En cruces de drenaje en corredores viales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (World Meteorological Organization, 2008; Moreno-Ponce et al., 2024).

Ejercicio 7.5. Modelación lluvia-escorrentía en carreteras

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con área, C, intensidad por duración y periodo de retorno. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Construya curvas IDF disponibles, seleccione duración próxima al tiempo de concentración y calcule caudal para varios periodos de retorno.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: tabla de caudales de diseño y sensibilidad al periodo de retorno. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

7.6 Simulación hidráulica de eventos extremos en infraestructura vial

La comprensión de simulación hidráulica de eventos extremos en infraestructura vial requiere separar el fenómeno observado de la forma en que se registra. En esta sección se integran periodo de retorno, obstrucción, socavación y capacidad residual; cada atributo conserva unidad, fecha, fuente y procedimiento de obtención. Esta decisión metodológica permite comparar

tramos y periodos sin convertir diferencias de medición en diferencias aparentes de seguridad. También obliga a distinguir entre un valor observado, un valor derivado mediante SIG o código y una predicción producida por el modelo (Chow et al., 1988; Dehnad & Yazdi, 2024).

Desde una perspectiva operativa, los escenarios hidráulicos extremos ayuda a estructurar la información y a identificar modos de falla antes de una emergencia. El análisis se plantea para obras expuestas a sedimentos y lluvias intensas, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (World Meteorological Organization, 2008; Chow, 1959).

7.6.1 Escenarios extremos y niveles de servicio hidráulico

En esta etapa, la fiabilidad de las mediciones resulta más importante que la complejidad inicial del algoritmo. Se emplean periodo de retorno, obstrucción, socavación y capacidad residual y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (World Meteorological Organization, 2008; Dehnad & Yazdi, 2024).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante los escenarios hidráulicos extremos se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. El caudal de diseño debe compararse con la capacidad hidráulica tanto en condición limpia como bajo escenarios razonables de obstrucción, sedimentación o aumento de rugosidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Chow et al., 1988; Moreno-Ponce et al., 2024).

7.6.2 Obstrucción parcial y pérdida de capacidad efectiva

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean periodo de retorno, obstrucción, socavación y capacidad residual y se revisa si cada medición corresponde al

mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

Para mantener la reproducibilidad, el análisis incorpora los escenarios hidráulicos extremos dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. El criterio de selección se define antes de observar el conjunto reservado para prueba (Dehnad & Yazdi, 2024; Federal Highway Administration, 2024).

La lectura ingenieril del resultado debe traducirse en una acción concreta: identificar modos de falla antes de una emergencia. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En obras expuestas a sedimentos y lluvias intensas, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Dehnad & Yazdi, 2024; Chow et al., 1988).

Para visualizar capacidad hidráulica frente al caudal de diseño y un escenario de obstrucción, la Figura 7.2 presenta una representación técnica que complementa la explicación y facilita su reproducción.

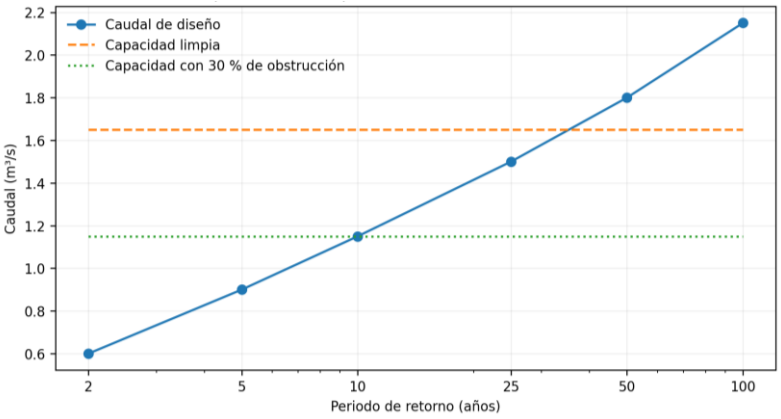


Figura 7.2. Capacidad hidráulica frente al caudal de diseño y un escenario de obstrucción.

Nota. Elaboración propia. Los caudales son ilustrativos y muestran el efecto de la pérdida de sección efectiva.

7.6.3 Socavación, erosión y modos de falla de la infraestructura

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran periodo de retorno, obstrucción, socavación y capacidad residual, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chow et al., 1988; Federal Highway Administration, 2024).

La evaluación utiliza los escenarios hidráulicos extremos, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Chow, 1959; Chow et al., 1988).

Los resultados deben interpretarse con una finalidad explícita: identificar modos de falla antes de una emergencia. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En obras expuestas a sedimentos y lluvias intensas, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Chow, 1959; World Meteorological Organization, 2008).

La estabilidad del resultado se evalúa mediante cambios controlados en los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (World Meteorological Organization, 2008; Moreno-Ponce et al., 2024).

7.6.4 Pruebas de estrés frente a lluvias de mayor intensidad

La evaluación se sustenta en periodo de retorno, obstrucción, socavación y capacidad residual, seleccionados por su relación directa con la decisión de ingeniería. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Chow et al., 1988; World Meteorological Organization, 2008).

La estimación utiliza escenarios hidráulicos extremos, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. El caudal de diseño debe compararse con la capacidad hidráulica tanto en condición limpia como bajo escenarios razonables de obstrucción, sedimentación o aumento de rugosidad. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Chow, 1959; Dehnad & Yazdi, 2024).

La simulación hidráulica aporta mayor valor cuando traduce el escenario meteorológico en una magnitud que pueda compararse con la capacidad de la obra. La Figura 7.3 muestra cómo cambia el caudal estimado cuando la intensidad de lluvia aumenta respecto del escenario base.

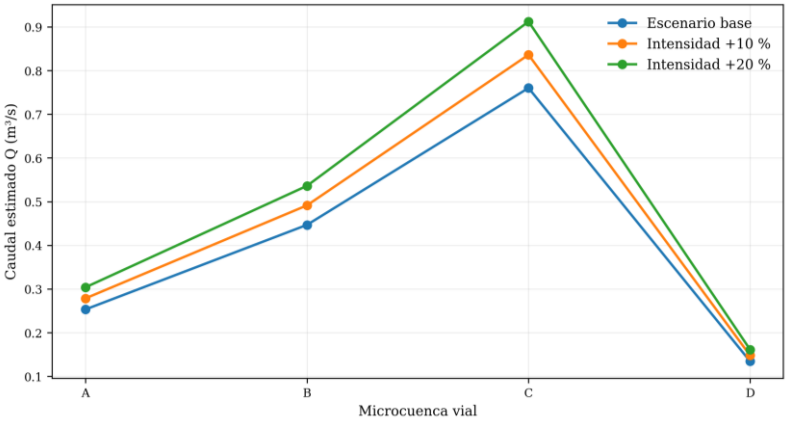


Figura 7.3. Escenarios de caudal estimado ante incremento de la intensidad de lluvia.

Nota. Elaboración propia a partir del Laboratorio Python 7; los incrementos de 10 % y 20 % son escenarios didácticos.

La separación entre escenarios permite reconocer obras con margen hidráulico reducido y orientar inspecciones de obstrucción, sedimentación o deterioro. El ejercicio no reemplaza un estudio hidrológico de diseño, pero muestra cómo incorporar sensibilidad climática y operacional a la priorización.

Ejercicio 7.6. Simulación hidráulica de eventos extremos

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con hidrograma o caudales de diseño, capacidad de obra y cota de calzada. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.

2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Simule escenarios de 10, 25, 50 y 100 años; calcule razón Qd/Qcap y marque los casos con excedencia o inundación de calzada.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: matriz de desempeño hidráulico por escenario. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

7.7 Integración de modelos hidrológicos con SIG para análisis de riesgo

Para trabajar con integración de modelos hidrológicos con sig para análisis de riesgo de forma reproducible, se establece primero la unidad de análisis y la pregunta que orienta el procedimiento. Se consideran DEM, rutas de flujo, inventario de drenaje y susceptibilidad, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Moreno-Ponce et al., 2024; World Meteorological Organization, 2008).

El procedimiento adoptado emplea SIG hidrológico y vincula cada salida con una acción verificable: priorizar mantenimiento y obras donde coinciden aportes y vulnerabilidad. La aplicación en Joa y corredores de la zona sur de Manabí obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (Federal Highway Administration, 2024; Chow, 1959).

7.7.1 Modelo digital de elevaciones y rutas preferentes de escorrentía

En este apartado, la evidencia se organiza a partir de DEM, rutas de flujo, inventario de drenaje y susceptibilidad. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Dehnad & Yazdi, 2024; Federal Highway Administration, 2024).

La validación exige contrastar la salida obtenida mediante el SIG hidrológico con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (World Meteorological Organization, 2008; Chow et al., 1988).

Los resultados deben interpretarse con una finalidad explícita: priorizar mantenimiento y obras donde coinciden aportes y vulnerabilidad. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En Joa y corredores de la zona sur de Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (World Meteorological Organization, 2008; Dehnad & Yazdi, 2024).

Antes de transferir el procedimiento a otro corredor se repite la evaluación con una partición espacial o temporal diferente. El ejercicio permite comprobar si la señal aprendida se conserva fuera del área original o si responde a particularidades locales. Esta evidencia se incorpora a la ficha de trazabilidad junto con versión de datos, parámetros, limitaciones y responsable de la revisión de campo (Moreno-Ponce et al., 2024; Federal Highway Administration, 2024).

7.7.2 Inventario georreferenciado de activos de drenaje

El apartado requiere una definición clara de las observaciones que se incorporan al análisis y de su forma de registro. Para esta parte se consideran DEM, rutas de flujo, inventario de drenaje y susceptibilidad, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Federal Highway Administration, 2024; Chow et al., 1988).

La estimación utiliza SIG hidrológico, con parámetros registrados y una línea base de comparación. Antes del ajuste se define cómo se separará la información destinada a calibración y comprobación; en modelos predictivos, esto incluye entrenamiento, validación y prueba. El caudal de diseño debe compararse con la capacidad hidráulica tanto en condición limpia como bajo escenarios razonables de obstrucción, sedimentación o aumento de rugosidad. La reproducibilidad requiere conservar la versión de los datos, los parámetros utilizados, la semilla aleatoria cuando corresponda y la fecha de ejecución (Moreno-Ponce et al., 2024; World Meteorological Organization, 2008).

7.7.3 Mapas de susceptibilidad a anegamiento y erosión

Para mantener la consistencia, se establecen de antemano las variables, unidades y reglas de registro. Para esta parte se consideran DEM, rutas de flujo, inventario de drenaje y susceptibilidad, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Chow et al., 1988; Chow, 1959).

La evaluación utiliza el SIG hidrológico, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La revisión de capacidad incluye condiciones de operación degradada, porque sedimentos, residuos o una mayor rugosidad pueden reducir de forma apreciable el caudal admisible. Todo cambio de

parámetro que modifique la prioridad de un tramo debe quedar documentado (Chow, 1959; World Meteorological Organization, 2008).

En términos de ingeniería, el análisis orienta la decisión de priorizar mantenimiento y obras donde coinciden aportes y vulnerabilidad. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En Joa y corredores de la zona sur de Manabí, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Chow, 1959; Dehnad & Yazdi, 2024).

Para ordenar capas sig útiles para riesgo hidráulico vial, la Tabla 7.2 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 7-2. Capas SIG útiles para riesgo hidráulico vial

Capa	Resolución/escala	Uso
DEM	m	Dirección de flujo
Red vial	línea	Activo expuesto
Drenajes	punto/línea	Capacidad e inventario
Suelos	polígono	Infiltración/erosión
Lluvia	estación/raster	Forzante hidrológica

Nota. Elaboración propia. Todas las capas deben compartir un sistema de referencia adecuado al análisis.

7.7.4 Priorización espacial del mantenimiento hidráulico

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de DEM, rutas de flujo, inventario de drenaje y susceptibilidad se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Moreno-Ponce et al., 2024; Dehnad & Yazdi, 2024).

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante el SIG hidrológico frente a una referencia ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. La capacidad disponible se verifica frente al caudal de diseño considerando el funcionamiento normal y escenarios plausibles de reducción por sedimentos u obstrucciones. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Dehnad & Yazdi, 2024; Federal Highway Administration, 2024).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: priorizar mantenimiento y obras donde coinciden aportes y vulnerabilidad. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En Joa y corredores de la zona sur de Manabí, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Dehnad & Yazdi, 2024; Chow, 1959).

Como verificación complementaria, se propone retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (Federal Highway Administration, 2024; World Meteorological Organization, 2008).

Ejercicio guiado 7.7. Integración de hidrología y SIG para análisis de riesgo

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con modelo digital de elevación, pendiente, drenaje, uso de suelo y red vial. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.

2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Reclasifique cada capa, normalice de 0 a 1, aplique una suma ponderada y cruce el resultado con segmentos viales.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: mapa de susceptibilidad hídrica y tramos que requieren comprobación de campo. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Laboratorio Python 7

El siguiente laboratorio implementa hidrología e hidráulica de cunetas. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 7. Método racional + capacidad de cuneta con Manning
# Requisitos: numpy, pandas
import numpy as np
import pandas as pd

# Escenarios didácticos de microcuencas viales.
data = pd. DataFrame({
 "microcuenca": ["A", "B", "C", "D"],
 "C": [0.55, 0.62, 0.70, 0.48],
 "I_mm_h": [92, 108, 126, 84],
 "A_ha": [1.8, 2.4, 3.1, 1.2]
})
# Q = 0.00278 C I A, con A en ha e I en mm/h.
data["Q_m3_s"] = 0.00278 * data. C * data. I_mm_h * data. A_ha
print(data)
```

```
# Cuneta trapezoidal: base b, tirante y, talud zH:1V.
def manning_trapezoid(b, y, z, n, S):
 A = y * (b + z*y)
 P = b + 2*y*np. sqrt(1+z*z)
 R = A/P
 Q = (1/n)*A*(R**(2/3))*np. sqrt(S)
 V = Q/A
 return Q, V

Qdis = data. Q_m3_s. max()
for y in np. arange(0.20, 1.01, 0.05):
 Qcap, V = manning_trapezoid(b=0.35, y=y, z=1.0, n=0.018, S=0.025)
 if Qcap >= 1.20*Qdis:
 print(f"Tirante adoptado={y:.2f} m; Qcap={Qcap:.3f} m3/s; V={V:.2f} m/s")
 break
```

Preguntas para el laboratorio

• ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?

• ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?

• ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?

• ¿Qué información de campo agregaría para reducir incertidumbre?

• ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

7.8 Referencias del capítulo 7

Federal Highway Administration. (2024). Urban drainage design manual (4th ed., Hydraulic Engineering Circular No. 22, FHWA-HIF-24-006).

Federal Highway Administration. (2012). Hydraulic design of highway culverts (3rd ed., HDS 5). U. S. Department of Transportation.

Chow, V. T. (1959). Open-channel hydraulics. McGraw-Hill.

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). Applied hydrology. McGraw-Hill.

World Meteorological Organization. (2008). Guide to hydrological practices (6th ed., Vols. I–II). WMO.

Dehnad, M. H., & Yazdi, A. (2024). A review of numerical and experimental studies on hydroplaning of vehicles in motion on road surfaces. Results in Engineering, 23, 102438. https://doi. org/10.1016/j. rineng.2024.102438

Moreno-Ponce, L. A., Álvarez-Álvarez, M. J., Delgado-Alvia, R. P., & Zavala-Vásquez, C. J. (2024). Importancia del coeficiente de fricción para el análisis de seguridad vial, en el tramo de carretera Jipijapa - La Pila de la provincia de Manabí. Revista Científica INGENIAR: Ingeniería, Tecnología e Investigación, 7(13, ed. esp.). https://doi. org/10.46296/ig. v7i13edespab.0188

Álvarez-Álvarez, M. J., Mejía Marcacuzco, J. A., Guevara Pérez, E., Velarde, E. C., & Regalado-Jalca, J. J. (2026). Integrated analysis of spatial water-quality gradients, hotspots, and inferred hydrological resilience using bioindicators and machine learning in a semi-arid river basin (Ecuador). Environments, 13(5), 278. https://doi. org/10.3390/environments13050278

Hernández Escobar, A. A., Fienco Bacusoy, A. R., Ruiz Palacios, J. P., & Pinargote Solorzano, N. J. (2025). Evaluación de susceptibilidad a deslizamientos en el Valle de Joa, cantón Jipijapa, provincia de Manabí. UNESUM-Ciencias. Revista Científica Multidisciplinaria, 9(2), 90-101. https://doi. org/10.47230/unesum-ciencias. v9. n2.2025.90-101

Modelación Predictiva y Análisis Espacial para la Seguridad Vial

Infraestructura en Territorios en Desarrollo

CAPÍTULO VIII

Infraestructura, Territorio y Modelación Predictiva para el Desarrollo

AUTOR: Orejuela Mendoza Ivanova Claribel

EDITORIAL RUNAIKI

8 Infraestructura, Territorio y Modelación Predictiva para el Desarrollo

Una vía modifica tiempos de acceso, costos logísticos, conexión con servicios y resiliencia de comunidades. En territorios con recursos limitados, la modelación predictiva debe priorizar inversiones que produzcan reducción de riesgo y mejoras de accesibilidad medibles. Se integran experiencias de Joa y Sancán para construir ejercicios reproducibles. En infraestructura rural, la accesibilidad se interpreta mejor mediante tiempos de viaje, continuidad anual y población servida. El Rural Access Index mide la proporción de población rural situada a menos de 2 km de una vía transitable todo el año, y los estudios de Joa y Sancán aportan evidencia territorial útil para ordenar inversiones sin convertirla indebidamente en información de siniestralidad (World Bank, 2024; Carvajal-Rivadeneira et al., 2025; Solórzano Villegas et al., 2026).

Resultados de aprendizaje

- Relacionar accesibilidad vial con producción, servicios y continuidad territorial en comunidades rurales.
- Construir indicadores de condición y necesidades de mantenimiento a partir de datos técnicos y comunitarios.
- Modelar conectividad mediante redes, centralidad, tiempos de viaje y Rural Access Index.
- Comparar alternativas de inversión con criterios técnicos, sociales, económicos y de riesgo.
- Diseñar una hoja de ruta de infraestructura que pueda actualizarse mediante indicadores verificables.

8.1 Relación entre infraestructura vial y desarrollo productivo

Antes de modelar relación entre infraestructura vial y desarrollo productivo, es necesario organizar la evidencia y comprobar que responda a la pregunta técnica planteada. La base reúne tiempo de viaje, confiabilidad, costo logístico y producción y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (World Bank, 2024; Cobos Lucio et al., 2025).

En términos operativos, la accesibilidad productiva facilita la organización de la evidencia necesaria para reconocer cómo la continuidad vial condiciona actividades locales. El análisis se plantea para Joa, Sancán y otras comunidades rurales, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Geurs & van Wee, 2004; Carvajal-Rivadeneira et al., 2025).

8.1.1 Tiempo de viaje y acceso efectivo a mercados

El análisis parte de una unidad de observación coherente con la carretera y con la finalidad del modelo. Los datos de tiempo de viaje, confiabilidad, costo logístico y producción se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Cobos Lucio et al., 2025; World Bank, 2024).

La interpretación requiere una comprobación externa; la salida obtenida mediante la accesibilidad productiva se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Regalado Jalca et al., 2025; Geurs & van Wee, 2004).

Los resultados aportan valor técnico cuando permiten reconocer cómo la continuidad vial condiciona actividades locales. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En Joa, Sancán y otras comunidades rurales, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Regalado Jalca et al., 2025; Solórzano Villegas et al., 2026).

8.1.2 Confiabilidad estacional de la vía para cadenas productivas

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean tiempo de viaje, confiabilidad, costo logístico y producción y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Cobos Lucio et al., 2025; Carvajal-Rivadeneira et al., 2025).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando la accesibilidad productiva como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Regalado Jalca et al., 2025; Cobos Lucio et al., 2025).

La interpretación se orienta a una decisión técnica definida: reconocer cómo la continuidad vial condiciona actividades locales. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En Joa, Sancán y otras comunidades rurales, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Regalado Jalca et al., 2025; World Bank, 2024).

Como control de sensibilidad, resulta útil retirar observaciones de gran influencia, variar la escala espacial o temporal y repetir el cálculo. El objetivo es determinar si el resultado depende de un grupo muy pequeño de datos o si expresa una relación estable. Las diferencias se informan junto con el modelo principal, porque la incertidumbre forma parte de la decisión y no debe ocultarse detrás de una única cifra (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

8.1.3 Costos logísticos y pérdidas por interrupciones

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean tiempo de viaje, confiabilidad, costo logístico y producción y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Geurs & van Wee, 2004; Carvajal-Rivadeneira et al., 2025).

La estimación utiliza accesibilidad productiva, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: reconocer cómo la continuidad vial condiciona actividades locales. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En Joa, Sancán y otras comunidades rurales, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (World Bank, 2024; Regalado Jalca et al., 2025).

8.1.4 Conectividad vial y oportunidades productivas de hogares rurales

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran tiempo de viaje, confiabilidad, costo logístico y producción, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin

depender de decisiones implícitas del analista (World Bank, 2024; Cobos Lucio et al., 2025).

La estimación utiliza accesibilidad productiva, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de reconocer cómo la continuidad vial condiciona actividades locales. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En Joa, Sancán y otras comunidades rurales, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Cobos Lucio et al., 2025; Carvajal-Rivadeneira et al., 2025).

Ejercicio 8.1. Infraestructura vial y desarrollo productivo

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con tiempo de viaje, condición vial, costo de transporte, producción y acceso a servicios. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Compare indicadores antes y después de una mejora o entre sectores con distinta calidad de vía; controle por distancia cuando sea posible.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: medida de beneficio territorial vinculada con una obra concreta. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

8.2 Diagnóstico técnico de infraestructura mediante análisis de datos

En diagnóstico técnico de infraestructura mediante análisis de datos, la primera decisión consiste en precisar qué representa cada observación y qué pregunta deberá resolver el análisis. Se consideran condición de caminos, puentes, drenaje y servicios, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Cobos Lucio et al., 2025; World Bank, 2024).

En los inventarios de infraestructura rural de Joa y Sancán, los índices técnicos y la evidencia comunitaria deben integrarse con criterios explícitos, datos versionados y una regla de decisión definida previamente. El análisis se orienta a convertir inspecciones dispersas en un diagnóstico comparable, mientras que la métrica funciona como criterio de apoyo y no como un fin en sí misma. La aplicación técnica requiere una comprobación externa mediante inspección, norma, medición independiente o información posterior al periodo de ajuste (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

8.2.1 Índices de condición para caminos, puentes y drenajes

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean condición de caminos, puentes, drenaje y servicios y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios

de codificación, precisión o cobertura entre campañas de levantamiento (Solórzano Villegas et al., 2026; Carvajal-Rivadeneira et al., 2025).

La interpretación requiere una comprobación externa; la salida obtenida mediante los índices técnicos y la evidencia comunitaria se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Carvajal-Rivadeneira et al., 2025; Regalado Jalca et al., 2025).

8.2.2 Integración de inspección técnica y percepción comunitaria

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de condición de caminos, puentes, drenaje y servicios se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante los índices técnicos y la evidencia comunitaria se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. El orden multicriterio depende de la normalización, los pesos y las restricciones; por ello, su estabilidad debe comprobarse antes de asignar presupuesto. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Regalado Jalca et al., 2025; Carvajal-Rivadeneira et al., 2025).

El resultado cuantitativo se convierte en una decisión operativa cuando permite convertir inspecciones dispersas en un diagnóstico comparable. Además de presentar resultados, el informe identifica la condición física que debe verificarse, el responsable de esa comprobación y los datos aún necesarios para pasar al diseño. En inventarios de infraestructura rural de Joa y Sancán, este paso permite asignar recursos de levantamiento a los sitios donde el costo de

una clasificación errónea sería mayor (Regalado Jalca et al., 2025; Cobos Lucio et al., 2025).

8.2.3 Datos mínimos para estimar necesidades de mantenimiento

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de condición de caminos, puentes, drenaje y servicios se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Cobos Lucio et al., 2025; Solórzano Villegas et al., 2026).

El resultado no debe aceptarse de forma mecánica. Por ello, la salida obtenida mediante los índices técnicos y la evidencia comunitaria se contrasta con una referencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. El orden multicriterio depende de la normalización, los pesos y las restricciones; por ello, su estabilidad debe comprobarse antes de asignar presupuesto. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Regalado Jalca et al., 2025; Cobos Lucio et al., 2025).

Para ordenar indicadores para diagnóstico de infraestructura rural, la Tabla 8.1 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 8-1. Indicadores para diagnóstico de infraestructura rural

Componente	Indicador	Evidencia
Vía	Condición superficial	Inspección y fotografías
Puente	Estado funcional	Ficha técnica
Drenaje	Obstrucción/capacidad	Levantamiento
Accesibilidad	Tiempo de viaje	Red y aforo
Servicio	Días de interrupción	Registro comunitario

Nota. Elaboración propia con orientación a diagnósticos de Joa y Sancán.

8.2.4 Jerarquización de deficiencias según servicio y riesgo

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean condición de caminos, puentes, drenaje y servicios y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Solórzano Villegas et al., 2026; Regalado Jalca et al., 2025).

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante los índices técnicos y la evidencia comunitaria frente a una referencia ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Carvajal-Rivadeneira et al., 2025; Regalado Jalca et al., 2025).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: convertir inspecciones dispersas en un diagnóstico comparable. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En inventarios de infraestructura rural de Joa y Sancán, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Carvajal-Rivadeneira et al., 2025; World Bank, 2024).

8.2.5 Aplicación territorial documentada: Joa y Sancán

La comuna Joa y el sitio Sancán permiten vincular el diagnóstico territorial con variables que pueden incorporarse a modelos de demanda, condición y priorización. En Joa, Zambrano Mero (2025) reportó 109 veh/día como tránsito actual y 310 veh/día para un horizonte de 20 años; el mismo estudio obtuvo un CBR crítico de 3,50 % y propuso una estructura de pavimento articulado de 8 cm de adoquín, 10 cm de base granular y 6 cm de subbase. Estos valores constituyen observaciones documentadas y deben conservarse separados de cualquier escenario generado posteriormente por el modelo.

Para Sancán, Carvajal-Rivadeneira et al. (2025) combinaron encuestas a 142 hogares con recorridos técnicos y reportaron que 61,9 % de los tramos evaluados se encontraba en mal estado, 32,4 % en condición regular y 5,7 % en buen estado. En movilidad, 72,5 % de los hogares indicó dificultades frecuentes de desplazamiento y 63,4 % reportó pérdidas o retrasos en actividades productivas asociados al estado de las vías. Estos indicadores permiten incorporar accesibilidad, condición física e impacto económico dentro de una misma lectura territorial sin confundirlos con una medición directa de siniestralidad.

Otro caso documentado en Sancán corresponde al camino Lomas del Jardín–Cecal. Loor Sánchez (2024) realizó aforos y diseño geométrico para el corredor; el estudio reporta un TPDA de 297 veh/día, clasificación C3 agrícola/forestal y velocidad de diseño de 40 km/h. Esta información resulta útil como línea base para escenarios de demanda y para analizar si la capacidad, geometría y programa de mantenimiento permanecen compatibles con el crecimiento futuro.

La Tabla 8.2 organiza los valores documentados que pueden emplearse como entrada o referencia de validación en los ejercicios predictivos del capítulo.

Tabla 8-2. Datos documentados de Joa y Sancán y su uso analítico

Territorio	Indicador	Valor documentado	Uso analítico
Joa	TPDA actual	109 veh/día	Línea base de demanda
Joa	TPDA a 20 años	310 veh/día	Referencia para proyección
Joa	CBR crítico	3,50 %	Condición de soporte del pavimento
Joa	Estructura propuesta	8 cm adoquín; 10 cm base; 6 cm subbase	Caso de diseño y mantenimiento
Sancán	Tramos en mal estado	61,9 %	Diagnóstico de condición vial

Sancán	Dificultad frecuente de desplazamiento	72,5 % de hogares	Indicador de accesibilidad
Sancán	Pérdidas o retrasos productivos	63,4 % de hogares	Impacto económico de la red
Sancán	TPDA Lomas del Jardín–Cecal	297 veh/día	Línea base de demanda
Sancán	Categoría y velocidad de diseño	C3 agrícola/forestal; 40 km/h	Referencia geométrica y operativa

Nota. Joa: Zambrano Mero (2025). Sancán: Carvajal-Rivadeneira et al. (2025) y Loor Sánchez (2024).

La distribución de los indicadores de Sancán facilita distinguir problemas de condición, movilidad y efecto productivo antes de construir un índice compuesto. La Figura 8.1 presenta una lectura gráfica de los resultados documentados.

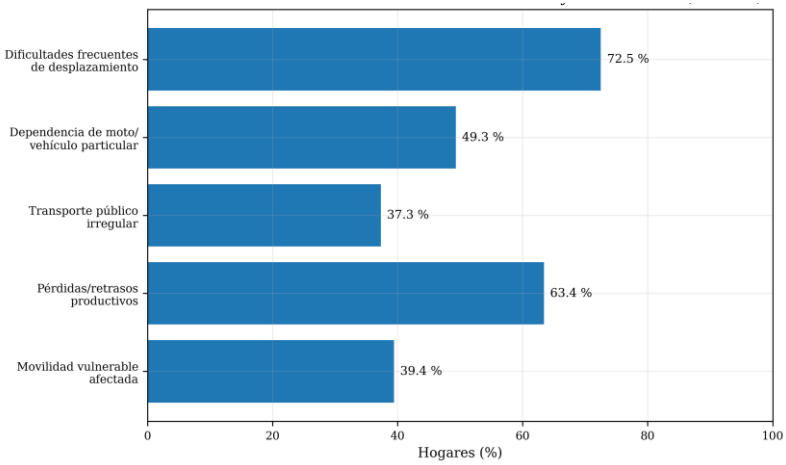


Figura 8.1. Indicadores documentados de movilidad y accesibilidad en Sancán.

Nota. Elaboración propia a partir de Carvajal-Rivadeneira et al. (2025), encuesta a 142 hogares.

Los porcentajes no representan probabilidades de accidente; describen condición vial, restricciones de movilidad e impactos productivos. Su valor para la modelación reside en que pueden incorporarse como indicadores de accesibilidad o vulnerabilidad territorial y contrastarse con inventarios físicos de la red.

Ejercicio 8.2. Diagnóstico técnico mediante análisis de datos

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con PCI o estado superficial, drenaje, señalización, TPDA, geometría y costo de mantenimiento. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Normalice variables técnicas y construya un índice de condición; verifique el resultado mediante inspección de tres segmentos.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: inventario de necesidades con evidencia física y costo preliminar. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

8.3 Modelación de accesibilidad y conectividad territorial

En modelación de accesibilidad y conectividad territorial, la primera decisión consiste en precisar qué representa cada observación y qué pregunta deberá resolver el análisis. Se consideran nodos, enlaces, tiempos, centralidad y cobertura, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el

mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Geurs & van Wee, 2004; Regalado Jalca et al., 2025).

En asentamientos rurales conectados con Jipijapa, el análisis de redes y el Rural Access Index deben aplicarse con parámetros documentados y datos versionados. El análisis se orienta a medir acceso a servicios con indicadores reproducibles, mientras que la métrica funciona como criterio de apoyo y no como un fin en sí misma. La aplicación técnica requiere una comprobación externa mediante inspección, norma, medición independiente o información posterior al periodo de ajuste (Carvajal-Rivadeneira et al., 2025; Geurs & van Wee, 2004).

8.3.1 Representación de la red mediante nodos y enlaces

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran nodos, enlaces, tiempos, centralidad y cobertura, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Regalado Jalca et al., 2025; Solórzano Villegas et al., 2026).

$$A_i = \sum_{j=1}^{m} O_j \, f(c_{ij})$$

El indicador A_i resume la accesibilidad del origen i a oportunidades O_j mediante una función de impedancia f(c_ij). El costo c_ij puede expresarse como tiempo, distancia o costo generalizado, según la finalidad del análisis (Geurs & van Wee, 2004).

El procedimiento se organiza desde la depuración hasta la evaluación independiente, utilizando análisis de redes y Rural Access Index como soporte analítico. Tras depurar la base, se obtienen las variables derivadas y se ajusta el modelo; la etapa de control revisa residuos, magnitud del error y estabilidad espacial. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. El criterio de selección se define antes de

observar el conjunto reservado para prueba (Solórzano Villegas et al., 2026; Cobos Lucio et al., 2025).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de medir acceso a servicios con indicadores reproducibles. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En asentamientos rurales conectados con Jipijapa, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Solórzano Villegas et al., 2026; World Bank, 2024).

8.3.2 Centralidad y conectividad de asentamientos y servicios

Al abordar el problema, la calidad de la evidencia pasa a ocupar el primer plano. Se emplean nodos, enlaces, tiempos, centralidad y cobertura y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Geurs & van Wee, 2004; Cobos Lucio et al., 2025).

La estimación utiliza análisis de redes y Rural Access Index, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (World Bank, 2024; Geurs & van Wee, 2004).

El procedimiento se integra a la práctica profesional cuando contribuye a la decisión de medir acceso a servicios con indicadores reproducibles. El informe debe responder a dos necesidades: facilitar la lectura de quien gestiona la red y permitir que otro profesional reconstruya el cálculo. En asentamientos rurales conectados con Jipijapa, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría

comprobar si la prioridad asignada tiene respaldo en terreno (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

Para visualizar red simplificada de accesibilidad entre asentamientos y servicios, la Figura 8.4 presenta una representación técnica que complementa la explicación y facilita su reproducción.

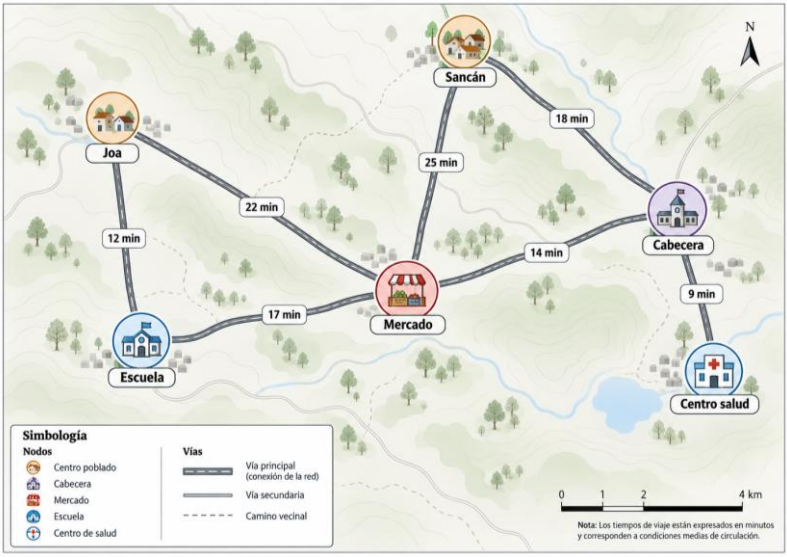


Figura 8.2. Red simplificada de accesibilidad entre asentamientos y servicios.

Nota. Elaboración propia. Los tiempos son ilustrativos y no representan mediciones oficiales de Joa o Sancán.

8.3.3 Rural Access Index y lectura territorial del acceso

El desarrollo del apartado requiere trabajar con nodos, enlaces, tiempos, centralidad y cobertura. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

El procedimiento combina análisis de redes y Rural Access Index con una verificación técnica que permita detectar resultados poco plausibles. La comparación sirve para detectar inconsistencias físicas, problemas de escala y mejoras marginales que no compensan un método más complejo. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Cobos Lucio et al., 2025; Solórzano Villegas et al., 2026).

Los resultados aportan valor técnico cuando permiten medir acceso a servicios con indicadores reproducibles. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En asentamientos rurales conectados con Jipijapa, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Cobos Lucio et al., 2025; World Bank, 2024).

8.3.4 Superficies de tiempo de viaje para evaluar cobertura

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran nodos, enlaces, tiempos, centralidad y cobertura, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (World Bank, 2024; Geurs & van Wee, 2004).

La estimación utiliza análisis de redes y Rural Access Index, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar información futura. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza

y la fecha en que se ejecutó el procedimiento (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

Los resultados aportan valor técnico cuando permiten medir acceso a servicios con indicadores reproducibles. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En asentamientos rurales conectados con Jipijapa, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Cobos Lucio et al., 2025; Solórzano Villegas et al., 2026).

Ejercicio 8.3. Accesibilidad y conectividad territorial

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con nodos de servicios, red vial, velocidades y población. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Calcule rutas mínimas y tiempos hacia salud, educación y mercados; compare accesibilidad actual con una mejora vial propuesta.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: mapa de tiempos y número de habitantes cuya accesibilidad mejora. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

8.4 Impacto de la red vial en sistemas económicos locales

Desde la ingeniería civil, el estudio de impacto de la red vial en sistemas económicos locales exige algo más que ejecutar una rutina estadística. El análisis combina costos de transporte, acceso a mercados, salud y educación y exige definir con anticipación qué representa cada observación, durante qué periodo es válida y qué incertidumbre acompaña a su medición. El modelo se construye después de esa revisión, de manera que el resultado conserve relación directa con la geometría, la operación o la infraestructura que posteriormente deberá inspeccionarse (Cobos Lucio et al., 2025; Carvajal-Rivadeneira et al., 2025).

En economías locales dependientes de la conectividad vial, las comparaciones territoriales y el análisis económico requieren criterios explícitos, datos versionados y reglas de decisión definidas antes de priorizar. La prioridad consiste en interpretar asociaciones sin atribuir causalidad automática. El desempeño estadístico se valora únicamente cuando mejora esa decisión. El resultado adquiere valor práctico cuando puede contrastarse con evidencia independiente, ya provenga de inspección, normativa, medición de campo o un periodo posterior (World Bank, 2024; Cobos Lucio et al., 2025).

8.4.1 Costos generalizados de transporte y actividad económica local

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de costos de transporte, acceso a mercados, salud y educación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Geurs & van Wee, 2004; Carvajal-Rivadeneira et al., 2025).

La estimación se obtiene mediante comparaciones territoriales y análisis económico y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. El orden multicriterio depende de la normalización, los pesos y las restricciones; por ello, su estabilidad debe comprobarse antes de asignar presupuesto. Todo cambio de parámetro que modifique la prioridad de

un tramo debe quedar documentado (World Bank, 2024; Solórzano Villegas et al., 2026).

La utilidad del análisis se hace evidente cuando permite interpretar asociaciones sin atribuir causalidad automática. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En economías locales dependientes de conectividad vial, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (World Bank, 2024; Cobos Lucio et al., 2025).

8.4.2 Acceso a educación y salud como beneficio territorial

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de costos de transporte, acceso a mercados, salud y educación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Geurs & van Wee, 2004; Carvajal-Rivadeneira et al., 2025).

La estimación se obtiene mediante comparaciones territoriales y análisis económico y se acompaña de una comprobación ingenieril independiente. El contraste entre resultados permite reconocer configuraciones poco plausibles, errores de escala o mejoras estadísticas demasiado pequeñas para justificar una mayor complejidad. El orden multicriterio depende de la normalización, los pesos y las restricciones; por ello, su estabilidad debe comprobarse antes de asignar presupuesto. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

La aplicación del procedimiento cobra sentido al apoyar la decisión de interpretar asociaciones sin atribuir causalidad automática. La salida técnica debe comunicar con claridad la decisión propuesta y conservar el detalle necesario para reproducir el procedimiento. En economías locales dependientes de conectividad vial, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que

permitiría comprobar si la prioridad asignada tiene respaldo en terreno (World Bank, 2024; Regalado Jalca et al., 2025).

8.4.3 Mercados agrícolas y sensibilidad a la confiabilidad vial

La evaluación comienza por delimitar una unidad de análisis que mantenga coherencia con la carretera y con la pregunta técnica. Los datos de costos de transporte, acceso a mercados, salud y educación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Solórzano Villegas et al., 2026; Regalado Jalca et al., 2025).

La secuencia metodológica parte de la depuración de los datos y culmina con una evaluación independiente; el análisis incorpora comparaciones territoriales y análisis económico. La secuencia continúa con el cálculo de variables derivadas, el ajuste del procedimiento y la revisión de residuos, error o estabilidad espacial, según corresponda. El orden multicriterio depende de la normalización, los pesos y las restricciones; por ello, su estabilidad debe comprobarse antes de asignar presupuesto. El criterio de selección se define antes de observar el conjunto reservado para prueba (Carvajal-Rivadeneira et al., 2025; Regalado Jalca et al., 2025).

8.4.4 Separación entre asociación económica y efecto causal

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de costos de transporte, acceso a mercados, salud y educación se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Solórzano Villegas et al., 2026; Regalado Jalca et al., 2025).

Para mantener la reproducibilidad, el análisis incorpora comparaciones territoriales y análisis económico dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La priorización multicriterio no es invariable: depende de los supuestos de normalización, ponderación y

restricción, que deben evaluarse mediante sensibilidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Carvajal-Rivadeneira et al., 2025; Regalado Jalca et al., 2025).

En términos de ingeniería, el análisis orienta la decisión de interpretar asociaciones sin atribuir causalidad automática. El resultado debe ser comprensible para la administración de la red y, al mismo tiempo, suficientemente detallado para que otro equipo pueda repetir el cálculo. En economías locales dependientes de conectividad vial, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Carvajal-Rivadeneira et al., 2025; World Bank, 2024).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (Regalado Jalca et al., 2025; Carvajal-Rivadeneira et al., 2025).

Ejercicio guiado 8.4. Red vial y sistemas económicos locales

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con tiempo de viaje, costos logísticos, ventas o producción, empleo y estado de la vía. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Construya una comparación temporal o entre zonas y separe asociación de causalidad; documente variables económicas que no se observan.

4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: estimación prudente del cambio económico asociado a la conectividad. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

8.5 Modelos predictivos para priorización de inversiones en infraestructura

Para analizar modelos predictivos para priorización de inversiones en infraestructura con rigor, primero se estructura la evidencia y luego se selecciona el procedimiento de modelación más apropiado. La base reúne condición futura, costo, población y beneficio esperado y conserva la procedencia de cada columna para que la cadena de cálculo pueda auditarse. Esta forma de trabajo evita dos errores frecuentes: entrenar con información que no habría estado disponible en el momento de la decisión y mezclar unidades territoriales que responden a procesos físicos diferentes (World Bank, 2024; Cobos Lucio et al., 2025).

En la práctica, modelos predictivos de priorización permite organizar la información para programar inversiones por etapas y con incertidumbre explícita. El análisis se plantea para carteras de obras comunitarias, donde la continuidad de los registros puede ser más importante que disponer de un inventario muy extenso levantado una sola vez. La evaluación compara error, estabilidad y facilidad de actualización antes de convertir un resultado numérico en una recomendación técnica (Geurs & van Wee, 2004; Carvajal-Rivadeneira et al., 2025).

8.5.1 Criterios de inversión y definición de variable objetivo

Para desarrollar el análisis, se adopta una unidad de observación compatible con la geometría vial y con la finalidad del modelo. Los datos de condición futura, costo, población y beneficio esperado se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si

representan un error o una condición vial real que precisamente interesa estudiar (Geurs & van Wee, 2004; World Bank, 2024).

La evaluación utiliza modelos predictivos de priorización, pero conserva una verificación ingenieril como referencia de plausibilidad. Este contraste permite comprobar si la mayor complejidad aporta una mejora real o solo introduce diferencias pequeñas sin utilidad ingenieril. La priorización multicriterio no es invariable: depende de los supuestos de normalización, ponderación y restricción, que deben evaluarse mediante sensibilidad. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

Los resultados deben interpretarse con una finalidad explícita: programar inversiones por etapas y con incertidumbre explícita. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En carteras de obras comunitarias, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (World Bank, 2024; Solórzano Villegas et al., 2026).

8.5.2 Pronóstico de deterioro y demanda de intervención

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean condición futura, costo, población y beneficio esperado y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Cobos Lucio et al., 2025; World Bank, 2024).

$$Q_t = Q_0(1 + r)^t$$

La expresión permite construir escenarios de crecimiento a partir de un valor inicial Q_0, una tasa anual r y un horizonte t. En el caso de Joa, cualquier tasa equivalente debe identificarse como escenario calculado y no como una serie histórica observada.

Para mantener la reproducibilidad, el análisis incorpora modelos predictivos de priorización dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La priorización multicriterio no es invariable: depende de los supuestos de normalización, ponderación y restricción, que deben evaluarse mediante sensibilidad. El criterio de selección se define antes de observar el conjunto reservado para prueba (Regalado Jalca et al., 2025; World Bank, 2024).

La lectura ingenieril del resultado debe traducirse en una acción concreta: programar inversiones por etapas y con incertidumbre explícita. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En carteras de obras comunitarias, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Regalado Jalca et al., 2025; Carvajal-Rivadeneira et al., 2025).

La verificación se completa mediante escenarios de sensibilidad y revisión profesional. Se observa qué ocurre cuando cambia la calidad del dato, la disponibilidad de una variable o el parámetro de mayor influencia. Una clasificación que solo se sostiene bajo una combinación muy específica debe tratarse con cautela; una que permanece estable ofrece una base más defendible para priorizar recursos de inspección (World Bank, 2024; Regalado Jalca et al., 2025).

Joa ofrece un ejemplo donde la predicción puede construirse a partir de dos valores publicados sin presentar como histórica una serie que no fue observada año por año. La Figura 8.3 reproduce la trayectoria equivalente entre 109 veh/día y 310 veh/día a 20 años e incorpora dos escenarios de sensibilidad para mostrar el efecto de modificar la tasa anual.

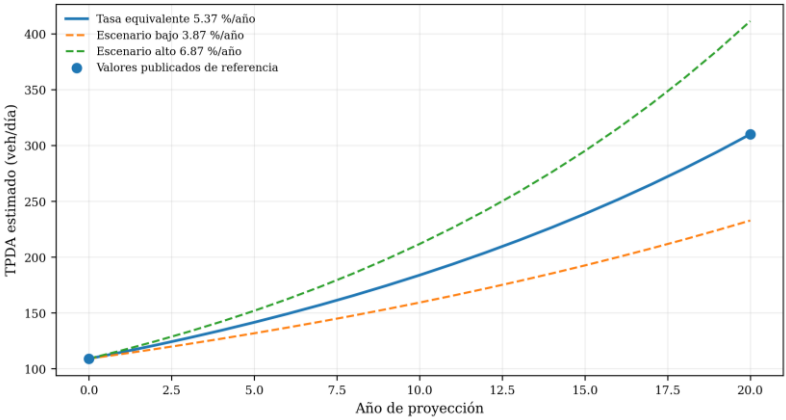


Figura 8.3. Proyección equivalente y sensibilidad del crecimiento vehicular en Joa.

Nota. Elaboración propia a partir de 109 veh/día y 310 veh/día a 20 años reportados por Zambrano Mero (2025). Los escenarios bajo y alto son didácticos.

La tasa equivalente del escenario central es aproximadamente 5,37 % anual. Este valor reproduce matemáticamente los extremos publicados, pero no debe interpretarse como una tasa histórica observada. La sensibilidad permite estimar cómo una variación razonable del crecimiento modificaría la demanda futura y, con ella, la programación de conservación o refuerzo estructural.

8.5.3 Beneficio esperado, costo y población atendida

En este apartado, la atención se centra en definir qué debe observarse y cómo debe registrarse. Para esta parte se consideran condición futura, costo, población y beneficio esperado, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (World Bank, 2024; Geurs & van Wee, 2004).

La estimación utiliza modelos predictivos de priorización, con parámetros registrados y una línea base de comparación. La estrategia de evaluación se establece antes de ajustar el procedimiento. Cuando existe aprendizaje estadístico, se separan entrenamiento, validación y prueba sin utilizar

información futura. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. Para repetir el análisis deben registrarse la versión de la base, los parámetros, la semilla aleatoria si se utiliza y la fecha en que se ejecutó el procedimiento (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

Los resultados aportan valor técnico cuando permiten programar inversiones por etapas y con incertidumbre explícita. Para ello se acompañan la estimación y su incertidumbre con evidencia de geometría, operación, condición o accesibilidad, según el problema estudiado. En carteras de obras comunitarias, una clasificación de prioridad sirve como filtro de inspección y no como autorización automática de construcción; el estudio de detalle conserva su papel profesional (Cobos Lucio et al., 2025; Solórzano Villegas et al., 2026).

Para ordenar variables para priorización predictiva de inversiones, la Tabla 8.3 reúne variables y criterios que pueden compararse sin separar el dato de su uso profesional.

Tabla 8-3. Variables para priorización predictiva de inversiones

Dimensión	Variable	Uso
Condición	Índice actual	Necesidad
Demanda	Usuarios/TPDA	Exposición
Costo	USD	Restricción
Beneficio	Tiempo/cobertura	Resultado esperado
Riesgo	Probabilidad de falla	Urgencia
Equidad	Población aislada	Distribución territorial

Nota. Elaboración propia. La priorización debe conservar los valores originales y los pesos usados en cada escenario.

8.5.4 Incertidumbre de datos y escenarios de programación plurianual

En este análisis, el desempeño del método depende primero de la calidad de los datos y solo después de la elección del algoritmo. Se emplean condición futura, costo, población y beneficio esperado y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y

temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Geurs & van Wee, 2004; Solórzano Villegas et al., 2026).

La estimación utiliza modelos predictivos de priorización, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (World Bank, 2024; Regalado Jalca et al., 2025).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de programar inversiones por etapas y con incertidumbre explícita. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En carteras de obras comunitarias, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (World Bank, 2024; Geurs & van Wee, 2004).

Ejercicio guiado 8.5. Predicción para priorización de inversiones

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con edad del pavimento, PCI, TPDA, drenaje, mantenimiento y costo de intervención. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.

3) Ejecute el método. Ajuste un modelo de deterioro o riesgo futuro y proyecte tres horizontes; convierta la predicción en año recomendado de intervención.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: programa plurianual de inversión con incertidumbre explícita. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Ejemplo breve. Ejemplo de decisión: una predicción de PCI=48 a dos años no ordena por sí sola reconstrucción. Debe combinarse con tránsito, seguridad, drenaje, costo y verificación de deterioros en campo.

8.6 Optimización de redes viales mediante modelos multicriterio

La aplicación de optimización de redes viales mediante modelos multicriterio requiere definir desde el inicio la unidad de análisis, el alcance temporal y la pregunta de ingeniería. Se consideran criterios, pesos, restricciones y alternativas, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

El procedimiento adoptado emplea MCDA y TOPSIS y vincula cada salida con una acción verificable: comparar proyectos con unidades y objetivos diferentes. La aplicación en programas de infraestructura con presupuesto limitado obliga a prestar atención a la disponibilidad real de datos, a la periodicidad del levantamiento y a la capacidad de la entidad responsable para actualizar el análisis. El desempeño se juzga fuera de la muestra utilizada para el ajuste y se acompaña con una lectura de incertidumbre y trazabilidad (World Bank, 2024; Solórzano Villegas et al., 2026).

8.6.1 Normalización de criterios con unidades heterogéneas

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran criterios, pesos, restricciones y alternativas, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (Regalado Jalca et al., 2025; Cobos Lucio et al., 2025).

$$z_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}$$

La normalización transforma cada criterio a una escala comparable. Para criterios de costo debe invertirse el sentido antes de aplicar la ponderación, de modo que todos los indicadores conserven una interpretación coherente.

La aplicación metodológica de MCDA y TOPSIS sigue una ruta ordenada: depuración, cálculo, ajuste y evaluación independiente. Las variables derivadas se calculan antes del ajuste, y la evaluación posterior examina residuos, error y estabilidad espacial de acuerdo con la técnica aplicada. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. El criterio de selección se define antes de observar el conjunto reservado para prueba (Solórzano Villegas et al., 2026; World Bank, 2024).

El análisis adquiere valor profesional cuando el resultado se traduce en una acción posible: comparar proyectos con unidades y objetivos diferentes. Ordenar valores es insuficiente. El documento técnico debe indicar qué condición física se comprobará, quién efectuará la revisión y qué información adicional se requiere antes del diseño. En programas de infraestructura con presupuesto limitado, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Solórzano Villegas et al., 2026; Carvajal-Rivadeneira et al., 2025).

8.6.2 Pesos técnicos y participación de actores

En este apartado, la unidad de análisis se define de manera que represente adecuadamente el tramo y la decisión de ingeniería. Los datos de criterios,

pesos, restricciones y alternativas se examinan mediante rangos físicos, consistencia entre campos y comparación con registros cercanos. Los valores atípicos no se eliminan de forma automática: primero se determina si representan un error o una condición vial real que precisamente interesa estudiar (Cobos Lucio et al., 2025; World Bank, 2024).

Para reducir el riesgo de sobreinterpretación, se verifica la salida obtenida mediante MCDA y TOPSIS frente a una referencia ajena al proceso de ajuste: medición de campo, norma, línea base o periodo posterior. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Regalado Jalca et al., 2025; World Bank, 2024).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de comparar proyectos con unidades y objetivos diferentes. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En programas de infraestructura con presupuesto limitado, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Regalado Jalca et al., 2025; Carvajal-Rivadeneira et al., 2025).

Para visualizar puntuaciones multicriterio ilustrativas para alternativas de infraestructura, la Figura 8.4 presenta una representación técnica que complementa la explicación y facilita su reproducción.

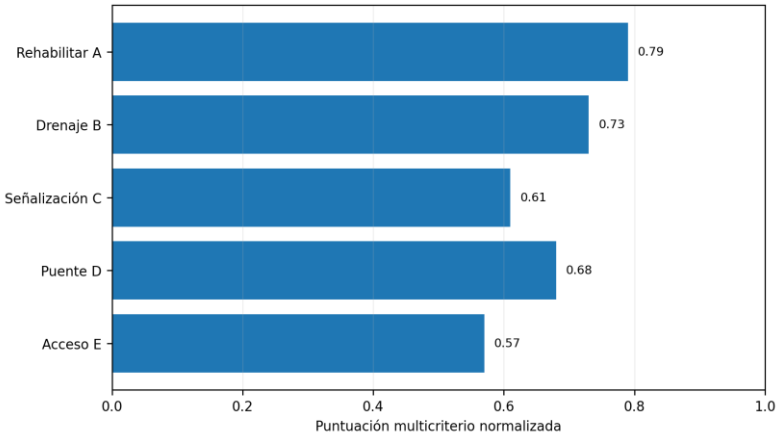


Figura 8.4. Puntuaciones multicriterio ilustrativas para alternativas de infraestructura.

Nota. Elaboración propia. Los valores son demostrativos; una aplicación real requiere criterios, pesos y costos verificados.

8.6.3 TOPSIS, ranking y frontera de alternativas factibles

Para abordar el problema se integran criterios, pesos, restricciones y alternativas, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Carvajal-Rivadeneira et al., 2025; Cobos Lucio et al., 2025).

La estimación utiliza MCDA y TOPSIS, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Geurs & van Wee, 2004; Carvajal-Rivadeneira et al., 2025).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de comparar proyectos con unidades y objetivos diferentes. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En programas de infraestructura con presupuesto limitado, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Geurs & van Wee, 2004; Solórzano Villegas et al., 2026).

8.6.4 Robustez de la selección ante restricciones presupuestarias

El análisis exige precisar las variables observables y la forma en que serán registradas. Para esta parte se consideran criterios, pesos, restricciones y alternativas, diferenciando los atributos que representan demanda, condición física, respuesta del sistema o exposición. Cada variable conserva definición operacional y unidad; si se deriva a partir de otra fuente, se registra la transformación utilizada. De esta forma, el análisis puede reconstruirse sin depender de decisiones implícitas del analista (World Bank, 2024; Cobos Lucio et al., 2025).

La estimación utiliza MCDA y TOPSIS, con parámetros registrados y una línea base de comparación. El ajuste comienza una vez definida la forma de evaluación independiente. En modelos predictivos se fijan previamente los conjuntos de entrenamiento, validación y prueba. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. La trazabilidad se mantiene documentando la versión de los datos, los parámetros del modelo, la semilla aleatoria cuando exista y la fecha de ejecución (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

Desde la gestión vial, el análisis sirve para fundamentar la decisión de comparar proyectos con unidades y objetivos diferentes. La comunicación del resultado debe ser clara para la gestión vial, sin perder la información metodológica necesaria para su reproducción. En programas de infraestructura con presupuesto limitado, se recomienda informar el valor central, un rango de incertidumbre, la fuente principal de error y la medida física que permitiría comprobar si la prioridad asignada tiene respaldo en terreno (Cobos Lucio et al., 2025; Carvajal-Rivadeneira et al., 2025).

Ejercicio guiado 8.6. Optimización multicriterio de redes viales

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con costo, beneficio, equidad, resiliencia, riesgo y factibilidad. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Aplique TOPSIS o AHP-TOPSIS, pruebe tres juegos de pesos y añada una restricción presupuestaria.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: cartera de proyectos seleccionada bajo presupuesto y análisis de sensibilidad. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

8.7 Lineamientos para planificación territorial basada en analítica avanzada

En lineamientos para planificación territorial basada en analítica avanzada, la primera decisión consiste en precisar qué representa cada observación y qué pregunta deberá resolver el análisis. Se consideran datos, responsables, escenarios, fases e indicadores, pero su incorporación no es automática: se revisan resolución, precisión, periodo de vigencia y compatibilidad con la variable objetivo. Cuando una fuente administrativa y un levantamiento de campo describen el mismo elemento con escalas distintas, la discrepancia se documenta antes de realizar cualquier ajuste (Cobos Lucio et al., 2025; World Bank, 2024).

La sección articula planificación territorial apoyada en analítica con una finalidad profesional concreta: mantener una cartera actualizable y verificable. En gestión interinstitucional de infraestructura rural, el procedimiento debe poder actualizarse con recursos razonables y dejar evidencia suficiente para que otro equipo repita el cálculo. Por ello se comparan alternativas simples y avanzadas mediante validación independiente; una diferencia pequeña en una métrica no compensa un método imposible de mantener, explicar o verificar en terreno (World Bank, 2024; Carvajal-Rivadeneira et al., 2025).

8.7.1 Gobernanza de datos territoriales y responsabilidades institucionales

El desarrollo del apartado requiere trabajar con datos, responsables, escenarios, fases e indicadores. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (Regalado Jalca et al., 2025; Cobos Lucio et al., 2025).

La interpretación requiere una comprobación externa; la salida obtenida mediante la planificación territorial apoyada en analítica se compara con una referencia que no participa en el ajuste: medición de campo, norma, línea base o periodo posterior. La posición de cada alternativa puede cambiar con los pesos, la normalización o las restricciones presupuestarias, de modo que el ranking requiere un análisis de sensibilidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Solórzano Villegas et al., 2026; Cobos Lucio et al., 2025).

La interpretación se orienta a una decisión técnica definida: mantener una cartera actualizable y verificable. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En gestión interinstitucional de infraestructura rural, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Solórzano Villegas et al., 2026; Geurs & van Wee, 2004).

La sensibilidad del resultado se estudia mediante variaciones controladas de los supuestos que más pueden influir en la clasificación. Se modifican parámetros, denominadores de exposición o reglas de agregación dentro de intervalos plausibles y se registra qué unidades cambian de posición. Si la prioridad es muy sensible, el resultado requiere información adicional; si permanece estable y coincide con evidencia independiente, puede utilizarse para orientar una inspección o una programación preliminar (Cobos Lucio et al., 2025; Solórzano Villegas et al., 2026).

8.7.2 Escenarios de crecimiento, riesgo y disponibilidad financiera

Para abordar el problema se integran datos, responsables, escenarios, fases e indicadores, procurando que cada variable conserve trazabilidad y utilidad ingenieril. El interés no está en reunir el mayor número posible de columnas, sino en conservar aquellas que puedan medirse, actualizarse y relacionarse con la decisión de ingeniería. Se documentan fuentes primarias y secundarias, resolución, fecha y cualquier regla de corrección aplicada a valores imposibles o incompletos (World Bank, 2024; Geurs & van Wee, 2004).

Para sostener la interpretación, el procedimiento basado en la planificación territorial apoyada en analítica se complementa con una comprobación directa desde la ingeniería. Al contrastar ambas salidas se pueden identificar resultados poco plausibles, errores de escala y aumentos de desempeño que carecen de relevancia práctica. Antes de utilizar el ranking para programar inversiones, se debe comprobar cuánto cambia al modificar pesos, escalas de normalización y restricciones. Todo cambio de parámetro que modifique la prioridad de un tramo debe quedar documentado (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

El resultado se lee en función de una finalidad de ingeniería claramente establecida: mantener una cartera actualizable y verificable. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En gestión interinstitucional de infraestructura rural, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Cobos Lucio et al., 2025; Solórzano Villegas et al., 2026).

Para Sancán, el TPDA publicado para el camino Lomas del Jardín–Cecal puede utilizarse como condición inicial de un análisis de sensibilidad. La Figura 8.5 compara tres tasas anuales hipotéticas de crecimiento y muestra cómo cambia la demanda esperada a diez años sin atribuir a ninguna de ellas carácter oficial.

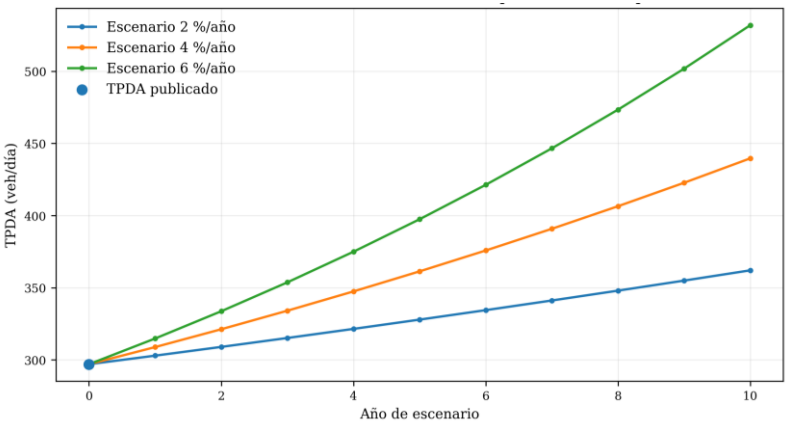


Figura 8.5. Escenarios didácticos de demanda a partir del TPDA publicado para Sancán.

Nota. Elaboración propia a partir de TPDA = 297 veh/día reportado por Loor Sánchez (2024). Las tasas de 2 %, 4 % y 6 % son escenarios de sensibilidad.

La comparación muestra que la incertidumbre sobre el crecimiento puede modificar de forma importante la demanda de diseño y el momento de intervención. En una aplicación real, las tasas deben estimarse con aforos históricos, cambios de uso del suelo y actividad productiva; el gráfico funciona como herramienta para probar decisiones frente a supuestos alternativos.

8.7.3 Hoja de ruta por fases para cartera de infraestructura

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean datos, responsables, escenarios, fases e indicadores y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre

campañas de levantamiento (Solórzano Villegas et al., 2026; World Bank, 2024).

La validación exige contrastar la salida obtenida mediante la planificación territorial apoyada en analítica con evidencia independiente del ajuste: medición de campo, norma, línea base o periodo posterior. La priorización multicriterio no es invariable: depende de los supuestos de normalización, ponderación y restricción, que deben evaluarse mediante sensibilidad. La evaluación se centra en desempeño, estabilidad y posibilidad de actualización, no solo en el valor obtenido durante el entrenamiento (Carvajal-Rivadeneira et al., 2025; Solórzano Villegas et al., 2026).

Los resultados deben interpretarse con una finalidad explícita: mantener una cartera actualizable y verificable. Esto requiere indicar qué tramo, activo o variable puede ser revisado y qué indicador deberá volver a medirse después de una intervención. En gestión interinstitucional de infraestructura rural, la recomendación conserva una separación estricta entre dato observado, estimación y medida propuesta; esa distinción impide presentar una predicción como si fuera evidencia de campo (Carvajal-Rivadeneira et al., 2025; Regalado Jalca et al., 2025).

8.7.4 Monitoreo de indicadores y actualización de prioridades

La evaluación comienza por comprobar la calidad de la evidencia disponible antes de ajustar cualquier modelo. Se emplean datos, responsables, escenarios, fases e indicadores y se revisa si cada medición corresponde al mismo tramo, periodo y escala. La consistencia espacial y temporal se comprueba antes de imputar, normalizar o transformar. Esta práctica evita que una aparente asociación nazca de cambios de codificación, precisión o cobertura entre campañas de levantamiento (Cobos Lucio et al., 2025; Geurs & van Wee, 2004).

Para mantener la reproducibilidad, el análisis incorpora la planificación territorial apoyada en analítica dentro de una secuencia que comienza con la depuración y termina con una evaluación independiente. Después del control inicial, el modelo se ajusta con las variables derivadas y se somete a una revisión de residuos, error y estabilidad espacial. La priorización multicriterio no es invariable: depende de los supuestos de normalización, ponderación y restricción, que deben evaluarse mediante sensibilidad. El criterio de selección

se define antes de observar el conjunto reservado para prueba (Regalado Jalca et al., 2025; Solórzano Villegas et al., 2026).

La lectura ingenieril del resultado debe traducirse en una acción concreta: mantener una cartera actualizable y verificable. El informe no debe limitarse a ordenar valores; también debe señalar la condición física que requiere comprobación, quién realizará la inspección y qué información falta antes de diseñar una obra. En gestión interinstitucional de infraestructura rural, este paso permite asignar recursos de levantamiento a los sitios donde el costo de una clasificación errónea sería mayor (Regalado Jalca et al., 2025; Cobos Lucio et al., 2025).

Ejercicio 8.7. Planificación territorial basada en analítica avanzada

Propósito. Aplicar el procedimiento de esta sección y dejar una cadena verificable desde el dato hasta la decisión de ingeniería.

Datos mínimos. Prepare 12 segmentos viales identificados S01-S12 con fuente, responsable, fecha de actualización, indicador, umbral y acción. Cada columna debe indicar unidad, fuente y fecha de medición.

1) Defina la unidad de análisis. Use segmentos de longitud homogénea o tramos delimitados por cambios geométricos. No mezcle en una misma fila dos sectores con condiciones físicas distintas.
2) Revise la base. Marque valores faltantes, duplicados y una observación imposible. Conserve el valor original en una columna y escriba la regla de corrección en otra; nunca sustituya un dato sin dejar registro.
3) Ejecute el método. Diseñe un ciclo de actualización de datos y modelos; registre versiones, responsables y criterios para activar inspección o inversión.
4) Compruebe el resultado. Compare con una línea base sencilla, una medición de campo, una norma técnica o una partición de prueba que no haya intervenido en el ajuste.
5) Traduzca el resultado. El producto mínimo será: protocolo institucional que permita repetir el análisis sin depender de una sola persona. No convierta una predicción en obra definitiva sin inspección y comprobación técnica.
6) 6) Entregue una ficha de una página con: datos usados, ecuación o algoritmo, parámetros, métrica, resultado por segmento, limitaciones y medida que sería revisada en campo.

Laboratorio Python 8

El siguiente laboratorio implementa topsis para priorización de inversiones. El código está escrito como un script completo: importa bibliotecas, crea o carga los datos incluidos, ajusta el método y produce una salida verificable. Puede copiarse a un archivo. py o ejecutarse por celdas. Para mantener reproducibilidad, los datos generados utilizan una semilla fija y los valores oficiales o locales se incorporan de forma explícita.

Entorno recomendado

Python 3.11 o superior. Bibliotecas principales: numpy, pandas, matplotlib, scipy, scikit-learn, statsmodels, geopandas, shapely, networkx y torch. Se recomienda crear un entorno virtual y registrar las versiones instaladas con `pip freeze`.

```
# Laboratorio 8. TOPSIS para priorización de intervenciones viales
# Requisitos: numpy, pandas
import numpy as np
import pandas as pd

# Matriz didáctica. Los valores locales publicados se incorporan donde están disponibles;
# otros campos son supuestos docentes que deben reemplazarse en una aplicación real.
df = pd. DataFrame({
 "alternativa": ["Joa-B", "Sancan-2", "Sancan-3", "Sancan-4", "Lomas-Cecal", "Jipijapa-Sancan"],
 "riesgo_seguridad": [7.2, 7.8, 6.9, 7.5, 8.0, 8.7],
 "personas_servidas": [5.5, 6.8, 6.1, 6.4, 7.1, 8.9],
 "deficiencia_drenaje": [6.2, 7.4, 6.8, 7.0, 7.6, 7.2],
 "costo_relativo": [4.8, 6.7, 6.1, 6.5, 5.7, 8.4],
 "factibilidad": [8.0, 6.5, 6.7, 6.4, 7.6, 7.0]
})
criteria = ["riesgo_seguridad","personas_servidas","deficiencia_drenaje",
 "costo_relativo","factibilidad"]
weights = np. array([0.32,0.22,0.18,0.16,0.12])
benefit = np. array([True, True, True, False, True])

X = df[criteria]. to_numpy(float)
R = X / np. sqrt((X**2). sum(axis=0))
V = R * weights
ideal_best = np. where(benefit, V. max(axis=0), V. min(axis=0))
ideal_worst = np. where(benefit, V. min(axis=0), V. max(axis=0))
d_plus = np. sqrt(((V - ideal_best)**2). sum(axis=1))
d_minus = np. sqrt(((V - ideal_worst)**2). sum(axis=1))
score = d_minus / (d_plus + d_minus)
df["topsis"] = score
df["prioridad"] = df["topsis"]. rank(ascending=False, method="min"). astype(int)
print(df[["alternativa","topsis","prioridad"]]. sort_values("prioridad"))
```

Preguntas para el laboratorio

- ¿Qué variable del ejercicio debe cambiarse primero si el objetivo se redefine?
- ¿Qué parte del código puede producir fuga de información si se aplica de manera incorrecta?

- ¿Qué métrica resulta más útil para una decisión de seguridad y por qué?
- ¿Qué información de campo agregaría para reducir incertidumbre?
- ¿Qué cambio permitiría transferir el ejercicio a otro cantón sin perder trazabilidad?

8.8 Referencias del capítulo 8

World Bank. (2024). Rural Access Index (RAI): Metadata and official data for SDG indicator 9.1.1. World Bank Data Catalog. https://datacatalog. worldbank. org/search/dataset/0038250/rural-access-index-rai

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

Roberts, P., Shyam, K. C., & Rastogi, C. (2006). Rural access index: A key development indicator. World Bank.

Carvajal-Rivadeneira, D. D., Cobos-Lucio, D. A., Cañarte-Baque, G. A., Zamora-Bailón, A. R., & Regalado-Jalca, J. J. (2025). Análisis de la infraestructura vial y su impacto en la movilidad y desarrollo económico en la comunidad de Sancán, cantón Jipijapa. Revista Científica INGENIAR: Ingeniería, Tecnología e Investigación, 8(15), 784–801. https://doi. org/10.46296/ig. v8i15.0271

Solórzano Villegas, L. E., Cordero Garcés, M. O., Cañarte Baque, G. A., Pin Morán, M. Y., Macías Cevallos, A. G., Carvajal Rivadeneira, D. D., Regalado Jalca, J. J., Indacochea Narváez, V. P., & Ponce Zavala, C. V. (2026). Infraestructura, territorio y recursos naturales: Diagnóstico, funcionalidad y gestión sostenible en comunidades rurales (Joa). RUNAIKI Editora-Editorial Internacional. ISBN 978-9942-571-35-9.

Regalado Jalca, J. J., Cobos Lucio, D. A., Carvajal Rivadeneira, D. D., Solórzano Villegas, L. E., Zavala Vásquez, C. J., Castro Chilán, D. M., Cedeño Molina, L. J., & Quizhpe Roman, J. M. (2025). Infraestructura y Desarrollo Comunitario en Áreas Rurales: Diagnóstico y Propuestas para el Sitio Sancán, Jipijapa. Runaiki Ediciones. ISBN 978-9942-571-10-6.

Cobos Lucio, D. A., Murillo Baque, D. S., Carvajal Rivadeneira, D. D., Solórzano Villegas, L. E., Domínguez Gálvez, D. L., Cordero Garcés, M. O., Castro Chilán, D. M., Reyes Calero, S. N., & Regalado Jalca, J. J. (2025). Pronóstico y Gestión de la Infraestructura Física:

Tendencias, Necesidades y Metas Estratégicas para el Sitio Sancán, Jipijapa, Ecuador. Runaiki Ediciones. ISBN 978-9942-571-15-1.

Loor Sánchez, O. D. (2024). Diseño geométrico de la vía que va desde Lomas del Jardín hasta el Cecal de la comuna Sancán del cantón Jipijapa [Trabajo de titulación, Universidad Estatal del Sur de Manabí]. Repositorio Digital UNESUM. http://repositorio. unesum. edu. ec/handle/53000/6525

Zambrano Mero, J. H. (2025). Diseño estructural para pavimento articulado en la calle B y sus intersecciones calles 6, 7 y 8 de la comuna Joa del cantón Jipijapa [Tesis de grado, Universidad Estatal del Sur de Manabí]. Repositorio Digital UNESUM. http://repositorio. unesum. edu. ec/handle/53000/7736

Ribera Carranza, E. S. (2025). Definición de la estructura del pavimento vehicular para las calles del sector 2, sitio Sancán, ciudad Jipijapa [Tesis de grado, Universidad Estatal del Sur de Manabí]. Repositorio Digital UNESUM. http://repositorio. unesum. edu. ec/handle/53000/7747

Gómez Santana, Á. J. (2025). Propuesta de implementación de medios de transporte para el sitio Sancán del cantón Jipijapa [Tesis de grado, Universidad Estatal del Sur de Manabí]. Repositorio Digital UNESUM. http://repositorio. unesum. edu. ec/handle/53000/8445

Álvarez-Álvarez, M. J., Mejía Marcacuzco, J. A., Guevara Pérez, E., Velarde, E. C., & Regalado-Jalca, J. J. (2026). Integrated analysis of spatial water-quality gradients, hotspots, and inferred hydrological resilience using bioindicators and machine learning in a semi-arid river basin (Ecuador). Environments, 13(5), 278. https://doi. org/10.3390/environments13050278