



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Comparación entre accidentalidad y accesibilidad a nivel urbano

Ing. Alejandro Londoño Vélez

Universidad Nacional de Colombia
Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Civil
Manizales, Colombia
2018

Comparación entre accidentalidad y accesibilidad a nivel urbano

Ing. Alejandro Londoño Vélez

Tesis presentada como requisito parcial para optar al título de:
Magister en Ingeniería – Infraestructura y sistemas de transporte
Línea de investigación

Director:

PhD Diego Alexander Escobar García

Línea de Investigación:

Accidentalidad vial y accesibilidad territorial

Universidad Nacional de Colombia
Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Civil
Manizales, Colombia

2018

A la negra, que me sentó a escribir.

*A los que apoyaron y a los que
observaron con duda.*

Agradecimientos

En primer lugar, se agradece la colaboración de la oficina de tránsito y transporte de la policía por haber facilitado la base de datos de accidentalidad usada en este proyecto.

Se reconoce y agradece el esfuerzo que hicieron las 3 estudiantes auxiliares del semillero de investigación de la oficina de movilidad, de la universidad, al ordenar las direcciones registradas en la base de datos de accidentes: Diana Becerra, Jessica Gil y Juliana Correa.

Se agradece la participación del tutor Diego Alexander Escobar en todas las etapas del proceso de investigación por los aportes de su experiencia y guía en los momentos de incertidumbre. De igual forma, se reconoce todo lo aportado en mi formación académica y laboral durante el tiempo compartido.

A los compañeros de la oficina de movilidad se agradecen las horas dedicadas a leer y comentar el presente trabajo para hacer de este un documento más completo.

Resumen

La accidentalidad vial es una de las mayores causas de muerte en el mundo, y actualmente es foco de esfuerzos de académicos y funcionarios gubernamentales. Al observar las cifras se tiene que Colombia no se queda atrás en esta problemática, presentando valores elevados con respecto a los estándares internacionales. Sin embargo, la toma de datos de accidentalidad en el país es muy deficiente y presenta falencias en la rigurosidad de los protocolos. Debido a esto se decide, en este trabajo, comprobar si la ocurrencia de accidentes se puede inferir a partir del análisis de la red vial presente en una ciudad.

Para esto se realiza un análisis de la accidentalidad de la ciudad de Manizales, apoyándose en los registros de accidentes viales de 10 años, usando el NAE, que es una medida ponderada que depende de la gravedad de los mismos, y se generan valores de densidad sobre el territorio con la estimación de densidad de núcleo (KDE). Estos valores se relacionan con los valores de tiempo medio obtenidos del procesamiento de la red vial e interpolados con la ayuda del modelo Kriging ordinario con semivariograma lineal.

A partir de los resultados obtenidos, se evidencia que las variables usadas presentan una correlación moderada pero debido a la escala en que se genera cada una, hacen imposible la inferencia de sitios particulares de accidentalidad.

Palabras clave: **Accidentalidad vial, accesibilidad, coeficientes de correlación.**

Abstract

Road accidents are one of the main death causes in the world, and currently, are the focus of efforts from academics and government officials. Looking at the statistics, Colombia has death rates greater than international goals. While overseeing the issue, the accident data collection in the country is very deficient and lacks protocol thoroughness. To tackle the issue of death rates, the objective of the thesis is to evaluate if the road accident occurrence can be inferred from the global mean accessibility of a city's road network.

In order to achieve this, an analysis of the Manizales' 10-year road accident data is performed using the EAN, a weighted measure based on the severity of the accident, and Kernel Density Estimation is used to interpolate values over the city. These resulting data is crossed with the mean time values obtained from the processing of the road network and interpolated using the Ordinary Kriging model with linear semivariogram.

From the results can be concluded that both variables present a moderate correlation but due to the scale in which each set of values is generated, an accurate inference of high road accident locations is impossible.

Keywords: Road accidents, accessibility, correlation coefficients.

Contenido

	Pág.
Resumen	IX
Abstract	X
Lista de gráficas	XII
Lista de figuras	XIII
Lista de tablas	XIII
Introducción	1
1. Antecedentes	5
1.1 Accidentalidad.....	5
1.1.1 En Colombia	8
1.1.2 Número de accidentes equivalentes	10
1.2 Accesibilidad	11
1.2.1 La red vial	14
1.2.2 Velocidad de la red	14
1.2.3 Kriging para la interpolación de datos	15
1.3 Correlación.....	16
1.3.1 Spearman	17
1.3.2 Kendall.....	18
1.3.3 Pearson	18
2. Metodología	21
2.1 Accidentalidad.....	22
2.1.1 Limpieza de la base de datos	22
2.1.2 Georreferenciación de los accidentes	24
2.1.3 Agregación de datos y visualización	25
2.2 Accesibilidad	27
2.2.1 Preparación de la red y asignación de velocidades	27
2.2.2 Cálculo de tiempos promedio.....	28
2.2.3 Interpolación	29
2.3 Correlación.....	29
2.3.1 Spearman	29

2.3.2 Kendall	30
2.3.3 Pearson	31
2.3.4 Coeficiente de determinación (R^2)	32
3. Resultados y análisis	33
3.1 Accidentalidad	34
3.2 Accesibilidad	42
3.3 Correlación	43
4. Conclusiones y recomendaciones	53
4.1 Conclusiones	53
4.2 Recomendaciones	54
Bibliografía	57

Lista de gráficas

	Pág.
Gráfica 1-1: Muertes violentas en Colombia, 2008-2017.....	10
Gráfica 2-1: Metodología de trabajo.....	22
Gráfica 2-2: Funciones de distribución de densidad.....	27
Gráfica 3-1: Densidad accidentes simples vs tiempo de viaje promedio	45
Gráfica 3-2: Densidad de NAE vs tiempo de viaje promedio	45
Gráfica 3-3: Densidad accidentes con daños vs tiempo de viaje promedio	47
Gráfica 3-4: Densidad accidentes con heridos vs tiempo de viaje promedio	47
Gráfica 3-5: Densidad accidentes con víctimas fatales vs tiempo de viaje promedio.....	48
Gráfica 3-6: DA con muertes y heridos vs tiempo de viaje promedio	50
Gráfica 3-7: DA con muertes y heridos con NAE vs tiempo de viaje promedio....	50

Lista de figuras

	Pág.
Figura 3-1: Accidentes en la ciudad de Manizales, 2007-2016.....	36
Figura 3-2: Mapa de calor de accidentes simples	37
Figura 3-3: Mapa de calor de accidentes con NAE	38
Figura 3-4: Densidad de accidentes con daños.....	39
Figura 3-5: Densidad de accidentes con heridos.....	39
Figura 3-6: Densidad de accidentes con víctimas fatales	40
Figura 3-7: Densidad heridos y muertos.....	41
Figura 3-8: Densidad heridos y muertos con NAE.....	41
Figura 3-9: <i>Raster</i> con tiempos de viaje promedio	43

Lista de tablas

	Pág.
Tabla 1-1: Muertes violentas y por accidentes viales en Colombia, 2008 - 2017... 9	9
Tabla 3-1: Distribución de accidentes por gravedad.....	35
Tabla 3-2: Coeficientes de correlación con distintos métodos y conjuntos de datos	51

Introducción

La accidentalidad vial es una de las principales causas de muertes violentas que padece la sociedad mundial desde hace algunos años, y en la actualidad es uno de los focos de esfuerzos tanto académicos como gubernamentales para ser contrarrestado con efectividad. En esta línea de acciones se congregan profesores, diseñadores viales y urbanísticos, y representantes gubernamentales para encontrar soluciones y campañas de mitigación a la mortalidad en accidentes viales.

Al adentrarse en las cifras de muertes se observa que en Colombia la accidentalidad no es la principal causa de muertes violentas, como sucede en el promedio mundial, sin embargo, no es debido a una baja accidentalidad sino a que los homicidios superan las estadísticas, pues en cuanto a la accidentalidad vial, el mundo tiene un promedio de 16.8 muertes por cada 100,000 habitantes, y Colombia presenta uno de alrededor de 13.3, que aunque por debajo del promedio, está muy alejado del 3 que presentan los países líderes del mundo en este tema.

El objetivo del presente trabajo es determinar si a partir de una medida de la red vial de una ciudad se puede inferir con algún grado de aproximación la ocurrencia de accidentes viales. Para determinar esto es necesario articular dos conceptos y sus derivadas medidas espaciales: la accidentalidad y la accesibilidad territorial.

Esta posibilidad de relación surge a raíz de la dificultad de encontrar datos históricos de accidentes viales en las ciudades colombianas que permitan, haciendo un análisis sencillo, una identificación de sectores de alto riesgo y ocurrencia de eventos. En este sentido, la evaluación de la red existente de la ciudad, derivando en un análisis de su accesibilidad territorial, puede resultar más

plausible en un municipio que no cuente con los registros de accidentes viales, y con base en estos resultados se podrían focalizar los esfuerzos por corregir las deficiencias que ocasionan los accidentes.

Como se decía, este análisis es posible al evaluar la accidentalidad y la accesibilidad de una ciudad, que en este caso será el municipio de Manizales, ubicado en Caldas, Colombia. Para la primera se tendrá en cuenta un registro histórico de 10 años de accidentes viales, y se evaluará la dispersión que estos tienen en el territorio, clasificándolos por la gravedad registrada, y a partir de esto generar un valor agregado usando 3 coeficientes de calificación, que producen el Número de Accidentes Equivalentes (NAE).

Para la accesibilidad territorial media global se tiene en cuenta la red vial con su respectiva velocidad de operación y se calculan los tiempos de viaje promedio para cada nodo de la red, lo que permite obtener una matriz de tiempos que describen las posibilidades de acceso al territorio a través de la infraestructura proporcionada. Los valores de un vector sumatoria de esta matriz se interpolan para generar valores sobre todo el territorio de la ciudad y poder relacionarlos con los valores de accidentalidad.

Las herramientas usadas para estos procesamientos y análisis son: software para la manipulación de bases de datos, sistemas de información geográfica (SIG), y coeficientes de correlación como el de Pearson, Kendall y Spearman.

Al realizar los cálculos de densidad sobre la base de datos de accidentes, se observa que estos se presentan en mayor cantidad sobre el centro y a lo largo de las avenidas principales de la ciudad, sectores que suelen presentar el menor tiempo promedio de accesibilidad, lo que coincide con la hipótesis de que a mejor accesibilidad mayor será el número y la gravedad de los accidentes.

Sin embargo, a pesar de que los coeficientes de correlación calculados muestran alguna relación, esta es apenas moderada, principalmente debido a que los accidentes presentan una dispersión menos suavizada que los tiempos promedio

de la red vial, y además, existe una proporción considerable de territorio que no presenta registro de accidentes y por lo tanto afecta el comportamiento combinado de las variables.

Los valores del coeficiente de correlación por rangos de Spearman corresponden a valores superiores de 0.5. Este resultado permite una interpretación de que si bien la ocurrencia de accidentes se relaciona con el tiempo medio de viaje de la red vial de la ciudad, esta puede deberse a muchos otros factores externos a la operación del tráfico en las vías, denominándose como un fenómeno multicausal.

En conclusión, se deduce que el análisis de accesibilidad media global de la red vial no permite una inferencia precisa de los lugares con alto índice de accidentalidad, principalmente debido a la diferencia en las escalas de afectación de cada variable.

1. Antecedentes

Al iniciar la investigación de este proyecto se indaga por la existencia de información académica que combinara los conceptos de accidentalidad y accesibilidad territorial, tanto cuantitativa como cualitativamente, pero los resultados de tal pesquisa fueron nulos. Se evidencia así que este acercamiento al tema no ha merecido la atención de la comunidad académica todavía, y debido a esto se genera un interés por explorarlo y entender qué relación pudieren llegar a tener los dos conceptos.

En este sentido, con el objetivo de abordar los temas de accidentalidad y accesibilidad de manera combinada, se debe inicialmente evaluarlos de forma separada, para entender las características principales que definen la accidentalidad como fenómeno social corregible y la accesibilidad territorial como medida de acceso del territorio dentro de la red vial disponible. Este capítulo se ocupa de explorar el contexto de ambos conceptos para entender su posible compatibilidad en un análisis combinado.

1.1 Accidentalidad

Desde un punto de vista global, al examinar las causas de muerte de las sociedades, se analizan tres grandes categorías: las muertes por enfermedades contagiosas, por enfermedades no contagiosas y por traumatismos o muertes violentas (World Health Organization, 2018a). Las dos primeras categorías contemplan todo tipo de enfermedades y condiciones de salud que atentan contra la vida, básicamente, en el largo plazo. La categoría de traumatismos o muertes violentas consta de eventos como el homicidio, el suicidio, muertes ocasionadas

por desastres naturales y los accidentes de tránsito que nos competen, entre otras.

A nivel mundial, la mayor causa de muerte dentro de esta categoría de traumatismos, son los accidentes de tránsito, seguidos por las caídas, los suicidios y posteriormente los homicidios (World Health Organization, 2018a). El tema de la accidentalidad vial en el mundo genera en la actualidad una preocupación latente, perceptible en los discursos de académicos y directivos estatales en las cumbres y congresos que se realizan sobre el tema (Bulc, 2017; Jones, 2016).

Los avances científicos y tecnológicos en áreas variadas de la salud, han hecho que las sociedades superen muchas de las enfermedades que antes las agobiaban. Asimismo, el refinamiento de la democracia y los procesos diplomáticos ha tenido como consecuencia la disminución de los conflictos armados de algunos países, y el cese, hasta ahora, de conflictos internacionales de escala mayor. Sin embargo, los líderes de la categoría de traumatismos demandan soluciones a la accidentalidad y campañas de salud mental y emocional para disminuir el número de suicidios.

Pensando en encontrar estas soluciones, las sociedades deben ponerse de acuerdo en el valor que puede tener una vida, y en este sentido, evaluar si es pertinente invertir en las soluciones que surgieran. Evidentemente, estos balances no se suelen hacer de manera explícita, aunque estimaciones del valor de una vida y de los heridos se han hecho en algunas comunidades de Europa (Niroomand & Jenkins, 2016), territorio que lleva la delantera en políticas de conservación social.

Si se le preguntase a una persona si mataría a otra por ahorrarse unos pesos, la respuesta seguramente sería negativa, pero los individuos solemos incurrir en aquellas acciones que indirecta o inconscientemente llevan a una muerte en el evento de un accidente, de allí su clasificación como tal.

Una acepción de accidente según la Real Academia Española (2001), es: “Suceso eventual o acción de que involuntariamente resulta daño para las personas o las cosas”. Un accidente es una situación imprevista y que por consiguiente altera el curso esperado de las cosas. La percepción de imprevisión de estos sucesos es la que se pretende cambiar, al identificar los comportamientos o deficiencias que ocasionan su ocurrencia, y de esta manera hacer responsable a algún sujeto en particular.

El Dr. Tedros Adhanom Ghebreyesus, director general de la Organización Mundial de la Salud desafía la definición de estos eventos viales como accidentes, cuando expresa que: “[...] *road traffic crashes are not “accidents”. They are completely preventable*” (World Health Organization, 2018b).

La perspectiva de algunos accidentes ha ido cambiando a medida que se investigan y socializan sus causas y métodos de prevención. La proliferación de esta información repercute en dejar de considerar un accidente como un evento fortuito, y entenderlo como una situación que se puede prevenir y mitigar. De este tipo de prácticas es que se han producido políticas públicas de prevención como lo es la exigencia de un certificado mecánico del vehículo, de condiciones de salud ideales del conductor, la utilización de elementos de protección mejor contruidos, etc.

Ahora bien, estos cambios en la normatividad gubernamental solo pueden desarrollarse desde la divulgación y masificación de la información de calidad, combinadas con una educación popular que sea capaz de hacer balances argumentativos con los resultados de las investigaciones que se produzcan. Es en este punto del proceso que entra esta investigación, como un insumo más a la gran discusión que es el desarrollo responsable de las comunidades modernas, que nos exige una madurez cada vez más alta en el desenvolvimiento en sociedad.

1.1.1 En Colombia

La accidentalidad vial en el mundo es considerada una de las mayores causas de mortalidad, particularmente la mayor en la categoría de traumatismos, como se ha dicho. En Colombia, 6754 personas murieron en 2017 debido a accidentes de tránsito, de las cuales más del 19.7% tenían entre 18 y 29 años, siendo el grupo etáreo con mayor proporción (Instituto Nacional de Medicina Legal y Ciencias Forenses, 2018). La corta edad de esta mayoría de víctimas, es de alta importancia para el costo social debido al potencial que tienen estas personas por desarrollar, ya que desde el punto de vista de la inversión social, en esta edad es cuando las personas han recibido una gran cantidad de recursos para su crecimiento y formación y apenas están comenzando a retribuir de vuelta esos recursos a la sociedad que se los brindó (World Health Organization, 2018b).

En la Tabla 1-1 se muestran los datos de muertes violentas en Colombia entre los años 2008 y 2017. Esta cifra de muertes violentas está compuesta por homicidios, muertes accidentales, suicidios, muertes indeterminadas y las muertes por accidentes de tránsito. En la tabla también se observan las muertes por accidentes viales y se calcula su proporción con respecto al total de muertes violentas para cada año. Adicionalmente, se presenta la tasa de muertes en accidentes de tránsito para cada año.

De estas tasas se puede observar que el promedio de esta década es de 13.35 muertes por cada 100,000 hab., lo que ubica a Colombia por debajo del promedio mundial que presenta un valor de 16.8, y asimismo, por debajo del promedio de la región de las Américas que presenta una media de 15.9 (Organización Panamericana de la Salud (OPS), 2016). Sin embargo, al analizar detalladamente los datos de los demás países, se observa como los menos mortales al Reino Unido o Suecia con tasas por debajo de 3.

Tabla 1-1: Muertes violentas y por accidentes viales en Colombia, 2008 - 2017.

Año	Muertes violentas	Muertes en accidente de tránsito	Proporción	Muertes en accidente de tránsito (Tasa/100.000 hab.)
2008	26958	5670	21.03%	12.76
2009	29435	5796	19.69%	12.89
2010	29962	5704	19.04%	12.53
2011	28996	5792	19.98%	12.58
2012	28496	6152	21.59%	13.21
2013	26538	6219	23.43%	13.20
2014	25225	6402	25.38%	13.43
2015	24681	6884	27.89%	14.28
2016	25438	7280	28.62%	14.93
2017	25381	6754	26.61%	13.70
Promedio	27111	6265	23.11%	13.35

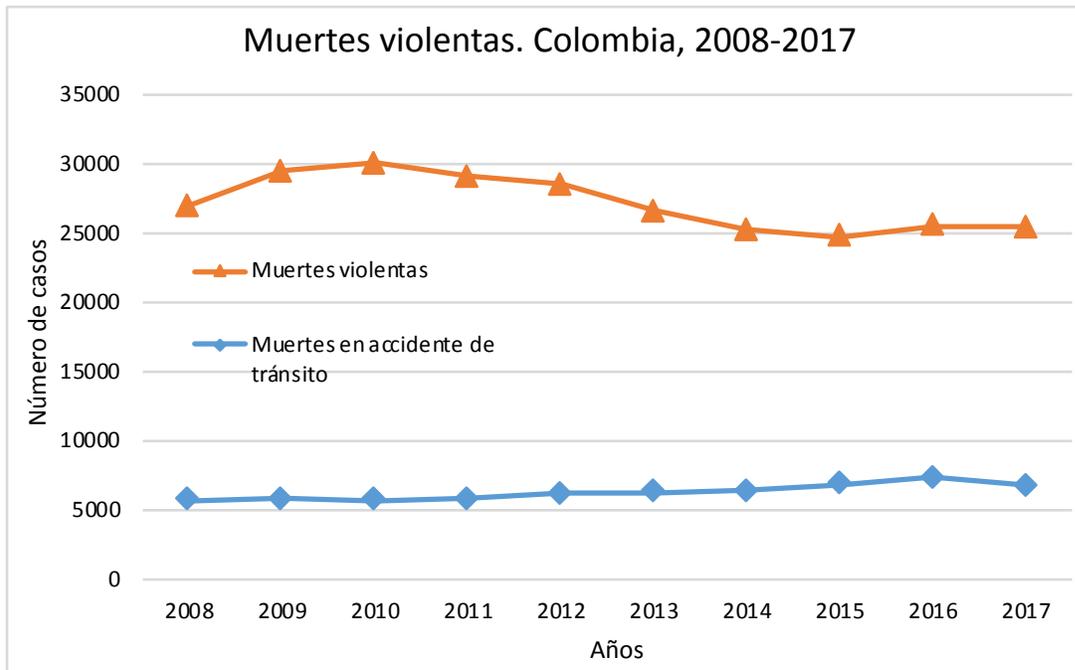
Fuente: Adaptado de (Instituto Nacional de Medicina Legal y Ciencias Forenses, 2018).

Los mejores reportes de estos países se deben a las políticas públicas que han implementado, el Reino Unido, por ejemplo, es famoso por la altísima rigurosidad con que se otorga la licencia de conducción, y Suecia, por su lado, es el país pionero en la visión cero de accidentalidad, que propone como meta la erradicación de muertes por este fenómeno (Tingvall & Haworth, 1999).

En la Gráfica 1-1 se observa cómo los casos de muertes violentas en Colombia han tenido una tendencia descendente en la última década. Por el contrario, cuando se presta atención a uno de sus componentes, como lo son las muertes en accidentes de tránsito, se identifica un aumento. Este fenómeno también es visible en la Tabla 1-1 al observar que la proporción de muertes que representan los accidentes viales ha ido aumentando con el paso del tiempo. Este comportamiento se debe a dos causas combinadas: el aumento en los casos fatales de accidentes de tránsito y la disminución de las muertes violentas en general, particularmente los homicidios.

En Colombia, los accidentes de tránsito son una de las mayores causas de mortalidad, superadas en la categoría de muertes violentas solamente por el homicidio, y seguidos por las muertes accidentales y el suicidio. Al comparar estas estadísticas con los datos del mundo, se observa que Colombia ha sido un país violento dado que los homicidios superan a los accidentes de tránsito, caso contrario en los datos a nivel mundial, sin embargo, estos últimos no pierden importancia en la estadística nacional, pues ocupan el segundo lugar en la lista, con 1 de cada 4 muertes violentas siendo por accidentes viales.

Gráfica 1-1: Muertes violentas en Colombia, 2008-2017.



Fuente: Adaptado de (Instituto Nacional de Medicina Legal y Ciencias Forenses, 2018).

1.1.2 Número de accidentes equivalentes

En el marco de las mediciones que se hacen en el tema de accidentalidad se hace necesaria la estimación de estos de forma numérica, para lo cual se utiliza un concepto que se conoce como Número de Accidentes Equivalentes (NAE, y en inglés Equivalent Accident Number, EAN), que se define como un número resultado de la ponderación de los accidentes de acuerdo con la gravedad de sus

consecuencias. Esto es, determinar si el accidente presentó víctimas fatales, heridos o solo daños a la propiedad material. La equivalencia del NAE se realiza con respecto a los accidentes de solo daños, con lo cual el factor multiplicante de estos siempre será 1.

Los valores de estos factores ponderantes pueden variar según el método que se use para estimar la equivalencia, incluso existen métodos donde se subdivide el grupo de heridos, en graves y leves, obteniéndose un total de 4 grupos para analizar (Sugiyanto, 2017). En estos casos se ha llegado a valores de 10: 4.25: 2.33: 1, para la ponderación de los accidentes en investigaciones realizadas en Tailandia (Sugiyanto et al., 2017). En México, por ejemplo, se han usado valores de 6:2:1 para calcular el NAE (Rascón Chávez et al., s. f.), teniendo en cuenta las tres categorías de gravedad principales.

En la ciudad de Manizales se ha usado la Ecuación (1.1) para estimar el valor del NAE con anterioridad en el Plan de Movilidad de Manizales del año 2011, y es con los valores de estos factores que se trabaja en el presente proyecto.

$$NAE = Acc.Fatal * 7.48 + Acc.Heridos * 1.43 + Acc.Simple * 1 \quad (1.1)$$

1.2 Accesibilidad

La accesibilidad del territorio, que es una oferta de la red de transporte (Añez et al., 1996), es causa de las dinámicas sociales que se generan en el espacio de una ciudad o del territorio en diferentes niveles (local, regional, nacional, continental) debido a la constante interacción entre los usos del suelo y los diferentes modos de transporte (Ben-Akiva & Lerman, 1979; Geurs & van Eck, 2001). Es decir, desde el punto de vista de desarrollo urbanístico, el potencial que tiene un espacio de ser alcanzado por la población define la manera en que ese entorno local es usado por la gente (Hansen, 1959) que utiliza la libertad de decisión para participar en las diferentes actividades ofertadas (Burns, 1980).

Este concepto posee registro de su aplicación desde las primeras décadas del siglo XX en temas relacionados con la planificación de sistemas de transportes en Norteamérica (Batty, 2009). Por otro lado, su definición clásica sugiere la capacidad de interacción que poseen los habitantes de una región teniendo en cuenta las impedancias de la red vial, entendidas en términos de distancia o tiempo hasta las oportunidades y el deseo de las personas de vencer dichas variables (Hansen, 1959). Más adelante, Dalvi & Martin (1976) definieron la accesibilidad teniendo en cuenta los diferentes modos de transporte disponibles para los habitantes y la facilidad de los mismos para llegar a las actividades ofrecidas en una ciudad mediante estos. Por su parte, Ingram (1971) tan solo la define teniendo en cuenta las oportunidades y sus diferentes variables de impedancia. Además, Geurs & van Eck (2001) tomaron conceptos de usos del suelo dado que su organización permite a los habitantes de una ciudad acceder a estos a través de diferentes modos de transporte. Asimismo, estos investigadores plantearon las perspectivas en las que se basa el concepto (infraestructura, equipamientos u oportunidades, personas y utilidades) y distintos componentes (usos del suelo, transporte temporal e individual) determinando la forma en que interactúan (Geurs & van Wee, 2004).

Por nuestra parte, la accesibilidad territorial se entenderá como el potencial de un lugar de ser alcanzado por los habitantes de su entorno dadas unas condiciones de viaje de la red de transporte (peatonal, transporte público, vehículo privado).

Al ser la accesibilidad una medida del potencial de una zona, permite intuir la dinámica social de esta, por ejemplo, el sector con mejor accesibilidad, suele ser lo que se llama el centro de la ciudad, que por regla general suele tener concentrados los servicios administrativos, financieros y demás actividades económicas que por extensión se ubican alrededor de estos. Zonas como esta van a atraer una gran cantidad de viajes, lo que puede significar un mayor riesgo de ocurrencia de un evento de accidente vial, relacionándose con la frecuencia de viajes y el nivel de exposición de las personas.

Ahora bien, hay varios niveles de accesibilidad en que se puede calcular la interacción de la red de infraestructura, y estas dependerán del interés de análisis en las posibles relaciones de desplazamiento (Ingram, 1971). De manera sencilla estas son:

- Accesibilidad relativa: tiene en cuenta la interacción de dos puntos de la red entre sí (Bunge, 1962).
- Accesibilidad integral: es la relación que existe entre todos los puntos de la red hacia uno solo específico. En otras palabras, es la agregación de muchos puntos con relación a uno de interés en el espacio.
- Accesibilidad global: es la relación entre todos los nodos de la red con el resto, o sea, el potencial de ir de un punto de la red a cualquier otro (Escobar & García, 2012; Talen & Anselin, 1998).

Usando este último nivel de accesibilidad se puede evaluar la condición general de la red e identificar el potencial de alguna intersección de interés. Este cálculo permite asignar un valor de accesibilidad al área de estudio. Este valor se da en unidades de tiempo, representando un tiempo bajo el mejor potencial de acceso al territorio. A partir de estos valores se pueden generar relaciones con otras medidas territoriales, como la contaminación esparcida en la ciudad (Escobar et al., 2016) o la accidentalidad que nos compete.

De esta manera se encuentra el potencial de uso de algunas zonas de convergencia de la ciudad, que típicamente son sitios que desarrollan una alta dinámica social, se vuelven espacios de encuentro para el intercambio cultural y económico, pero al mismo tiempo son lugares de paso, lo que ocasiona diferencias en el modo en que se usan y debido a esto, entre los usuarios se generan conflictos, situación que propicia accidentes. Debido a estas premisas surge la hipótesis de que a mayor accesibilidad, mayor el número de accidentes que ocurren en un sector.

1.2.1 La red vial

En la actualidad se da por sentada la existencia de un archivo de la red de transporte de algún territorio debido al avance que existe en la generación de inventarios. Sin embargo, al momento de emprender un proyecto que deba usar este insumo informático, se debe verificar como mínimo su continuidad y conectividad, es decir, que al ser una red, cualquier cálculo que se realice sobre ella dependerá de las conexiones que presente. Además, dado que en este caso se pretende obtener valores de un nodo a todo el resto de nodos que constituyen la red, esta deberá ser cerrada, esto es, que no podrá presentar agrupamientos separados, *clusters*, o cúmulos independientes, es decir, que todos los arcos deben estar conectados, haciéndose posible la llegada a todos los otros nodos a partir de cualquiera.

1.2.2 Velocidad de la red

Una de las variables más importantes en el cálculo de la accesibilidad es la velocidad de la red. Algunos acercamientos se han hecho para estimarla con bases teóricas, yendo de lo más simple como tener en cuenta el perfil vial de cada arco, dividiéndolo en varios niveles de red y asignando una velocidad a cada uno, o métodos complejos como simular el tráfico de la red y obtener velocidades promedio de los segmentos de esta.

De otro lado, también se han explorado varios métodos empíricos para medir tanto la velocidad como el flujo de la red, incluyendo procesos de medición fijos y móviles, tales como conteos manuales, sensores piezoeléctricos, detección de imagen de video, monitoreo de redes móviles y recolección de datos de GPS (Leduc, 2008; Wang et al., 2012).

En este sentido, mientras los modelos teóricos se refinan, las mediciones de campo son una buena solución para la adquisición de esta información, siempre y cuando existan recursos para ello.

1.2.3 Kriging para la interpolación de datos

A partir de la red vial con la información de velocidad, se da paso al cálculo de los tiempos promedio de cada nodo constituyente de la red. Con el objetivo de obtener la interpolación de estos valores de tiempo distribuidos en el espacio, se hace uso del modelo geoestadístico Kriging Ordinario, que permite determinar la estructura de autocorrelación de los datos, usando métodos que minimizan el error cuadrático medio de la predicción.

Entendiendo la importancia del semivariograma para la generación de un modelo, que estime los valores interpolados reflejando una correlación espacial de los datos contenidos en los nodos, se elige trabajar con un semivariograma lineal.

El modelo de Kriging usa un principio, como la mayoría de interpolaciones geoespaciales, que se conoce como la Primera Ley de Tobler de la Geografía (Tobler's First Law of Geography), que dice que "Todas las cosas están relacionadas entre sí, pero las cosas más próximas en el espacio tienen una relación mayor que las distantes" (Tobler, 1970). Esta va de la mano con el principio de autocorrelación espacial (Haining, 2015).

Algunos cuestionamientos se han hecho a la Ley de Tobler (Waters, 2018), poniendo en discusión si es en realidad una ley y a qué campos específicos de análisis geoespacial puede o no aplicar. Particularmente se cuestiona su pertinencia y validez sobre información que surge de las redes, como la malla vial de una ciudad.

Esto se debe a que en estas redes, pueden existir valores cercanos en el espacio, pero distantes en la red, como es el caso de una doble calzada, en que un sentido de esta puede presentar atributos distintos al otro sentido y sin embargo estar separadas solo unos pocos metros en el espacio.

1.3 Correlación

Una de las metodologías estadísticas más usadas en las investigaciones es el análisis de correlación. Con esta es posible determinar el grado de asociación lineal entre dos o más variables estudiadas. Sin embargo, la relación existente entre las variables no es necesariamente por razón de causalidad, advirtiéndose el cuidado que se debe tener al hacer inferencias de este tipo. Cuando esta relación se puede explicar por la presencia de una tercera variable, se le denomina correlación espuria (Haig, 2004).

Determinar la relación entre variables al realizar análisis de datos es importante para las investigaciones porque permite establecer de forma más clara las interacciones que tienen distintos aspectos de un fenómeno, que puede ir de lo físico a lo sociológico. Las medidas de asociación calculadas son medidas estadísticas descriptivas que reflejan dirección y fuerza de relación entre dos o más variables, y no se deben confundir con pruebas estadísticas inferenciales (Sheskin, 2000).

Debido a que este procedimiento pretende examinar la relación entre las diferencias individuales, los datos obtenidos deben ser pares de observaciones, con datos de ambas variables para un sujeto o caso. De esta manera, la medida de correlación muestra la intensidad de relación entre las variables al determinar la cercanía de los pares de coordenadas con una línea recta, ya sea de pendiente positiva o negativa. Una relación fuerte tendrá el conjunto de datos muy cerca de la línea, de lo contrario, con una alta dispersión de estos, se dice que la relación es débil.

En cuanto a la relación que tienen los términos *correlación* y *causalidad*, se puede decir que el primero se refiere únicamente al estudio de la magnitud y sentido de la relación entre variables, constatando la existencia de una asociación lineal entre ambas, mientras el segundo considera la causa y efecto de un fenómeno, que no siempre están relacionadas en sus valores medibles, o que se relacionan con el desfase de alguna dimensión, normalmente el tiempo.

En este sentido, cuando parece que las dos variables estudiadas están relacionadas, no se debe necesariamente a que una está causando la otra, puede ser que la relación sea casual o que la verdadera causa de ambas variables sea una tercera variable, lo que se traduce en una relación espuria de las dos primeras (Kenny, 1979).

La correlación monótona entre variables se puede estimar usando tres de los métodos más populares. Estos son los coeficientes de Spearman, Kendall y Pearson.

1.3.1 Spearman

El coeficiente de correlación por rangos de Spearman resulta muy útil cuando las variables de estudio no presentan un comportamiento normal. Este coeficiente se puede entender como el mismo coeficiente de Pearson pero aplicado a la transformación de las variables en rangos clasificados (Pardo Merino & Ruiz Díaz, 2002).

El grado de asociación lineal obtenido de este coeficiente no corresponde a los valores de las variables observadas sino a las clasificaciones por rangos de estas. Esto significa que si el coeficiente resulta en un valor alto, no necesariamente significa que los datos sigan la forma de una línea, puesto que este método es independiente de la forma que pueda tener la relación, solo se basa en el sentido que siguen sus cambios de valor.

El coeficiente se basa en las concordancias y discordancias existentes entre las clasificaciones o posicionamientos de los valores de las variables. El resultado variará de -1 a +1, tomando estos valores extremos cuando la relación de las variables sea alta, y el signo definirá el sentido de la relación. Valores cercanos a 0 significan una muy baja relación de las observaciones analizadas.

Esta medida estadística es muy sensible a los valores atípicos en las observaciones. Con la presencia de registros que se salgan de manera extrema

del comportamiento general de los datos, el coeficiente reduce su valor de forma drástica.

1.3.2 Kendall

El coeficiente de correlación de Kendall se recomienda para variables cualitativas de tipo ordinal, o bien cuantitativas que han sido ordenadas por su medida. La aplicación de este coeficiente es pertinente cuando las variables de estudio no presentan una distribución poblacional conjunta normal (Morales & Rodríguez, 2016).

Esta medida de correlación suele presentar valores menores que el coeficiente de Spearman, particularmente cuando los datos presentan pocos valores atípicos (extremos).

La formulación de este coeficiente permite una variación para tener en cuenta la presencia de empates en la ordenación de los datos de la muestra. Esto es beneficioso cuando se sabe que los datos son cuantitativos y la muestra va a ser grande, con lo que se aumenta la posibilidad de tener registros repetidos en las variables estudiadas.

1.3.3 Pearson

El coeficiente de correlación de Pearson es una medida estadística aplicable a variables cuantitativas y continuas y es usado para estimar la fuerza de dependencia lineal entre estas. Se diferencia de la covarianza en tanto que es independiente de la escala de medida de las variables, lo que lo hace adecuado para realizar comparaciones entre comportamientos de distintas variables.

El coeficiente de Pearson puede tomar valores entre -1 y +1, siendo el signo el sentido de la relación entre las variables y la magnitud el grado de asociación lineal que tengan. Cuando el coeficiente da como resultado valores muy cercanos a 0, se dice que la relación lineal entre las variables es muy débil o nula, pero

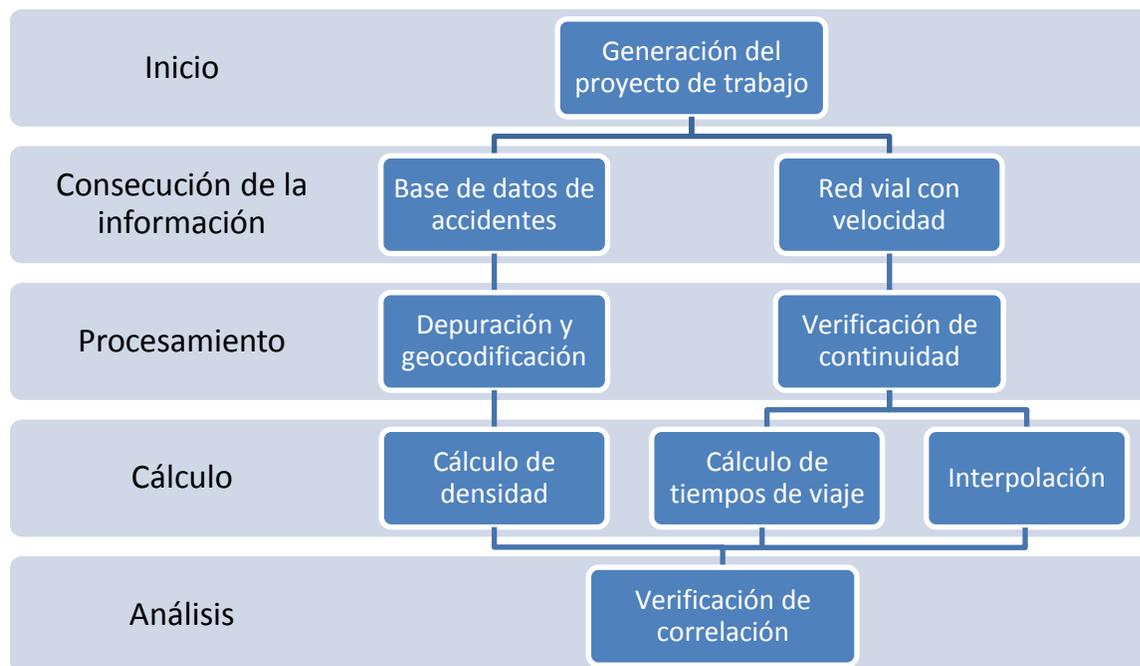
esto no quiere decir que necesariamente las variables son independientes, pues puede existir una relación que responda a una forma distinta.

2. Metodología

En este capítulo se explican los detalles de los procesos de cálculo que se realizan para el desarrollo de la tesis, entre los cuales está el procesamiento de datos, la georreferenciación y las teorías tenidas en cuenta para el análisis de los resultados obtenidos.

Se explicarán primero los detalles metodológicos para el procesamiento de los datos de accidentalidad utilizados en el caso de estudio, luego se mostrarán las condiciones del modelo para el cálculo de la accesibilidad territorial usada como insumo y por último se detallará cómo se combinaron los datos de ambos campos de estudio para obtener los resultados pertinentes de análisis al final del trabajo.

En la Gráfica 2-1 se muestra el diagrama básico de la metodología usada. En esta se observan las etapas de consecución de información, procesamiento y cálculo que se realizan para cada concepto por separado, y en última instancia el análisis del resultado de combinar ambas variables a través de varios coeficientes de correlación.

Gráfica 2-1: Metodología de trabajo.

Fuente: Elaboración propia

2.1 Accidentalidad

En el apartado de accidentalidad el desarrollo de la metodología que se sigue es:

La consecución de la base de datos, la depuración y estandarización de la información, procesarla con herramientas que permitan la geocodificación y así automatizar la obtención de coordenadas geográficas de cada registro, generar la representación gráfica de estas coordenadas sobre el mapa de la red vial de la ciudad, crear a partir de esta dispersión la estimación sobre el territorio de la densidad de incidentes viales, y por último relacionar estos datos con los tiempos de accesibilidad territorial de la red.

2.1.1 Limpieza de la base de datos

Al iniciar, la base de datos suministrada por la secretaría de tránsito del municipio se encontraba en un archivo de hoja de cálculo a manera de base de datos pero

con inconsistencias de formato, debido a que no se había limpiado en los años en que se había estado llenando de manera diligente.

Para corregir esta falencia de la información, que resulta muy común en países en desarrollo, debido a la falta de exigencia de protocolos estandarizados que permitan una mejoría en la calidad de los procesos, en primer lugar se debía estandarizar, particularmente en el registro de la dirección, que es lo que posteriormente se usa para geocodificar cada registro, obteniendo sus coordenadas geográficas. Previo a la limpieza en lotes de la información, se debe separar las partes de la dirección. Este proceso se realizó con la ayuda de varias estudiantes voluntarias al proyecto, obteniéndose la dirección registrada separada en sus componentes alfanuméricos ordenados.

Al tenerse la información ordenada de esta manera se puede proceder a limpiarla con la ayuda del software OpenRefine, que es una herramienta que permite explorar, limpiar y enlazar con diferentes servicios web la información que se tenga, permitiendo de manera sencilla y eficiente para grandes bases de datos, cambiar palabras que aparecen escritas de forma distinta para estandarizar la forma en que se debe tratar los datos.

Por ejemplo, cuando se registra la dirección de los accidentes por parte de la autoridad de tránsito, se incurre en diferencias al momento de nombrar o abreviar los nombres de las calles de la ciudad, como en la abreviatura de Avenida, que se usa como “Av.”, “avn” o “av”, lo que genera dificultades al suministrar los valores a las herramientas de geocodificación a utilizar.

Para lograr la limpieza total de la información, se deben analizar las características de esta y usar tanto el lenguaje de programación Python, integrado en el funcionamiento de OpenRefine, como el lenguaje GREL, propio del software que era mantenido con anterioridad por Google, de lo que se deriva su nombre (Google Refine Expression Language). Ambos lenguajes de programación no requieren niveles avanzados de conocimiento en sistemas para

lograr las tareas de edición de texto y separación o concatenación de columnas en la base de datos, lo que hace de este procedimiento una tarea sencilla de realizar para los ingenieros de tránsito que pretendan hacerlo.

Finalmente, al tener estandarizada y homogénea la información de la dirección en la base de datos se puede proceder a tomarla como insumo para la geocodificación automática de cada registro.

2.1.2 Georreferenciación de los accidentes

Con el objetivo de relacionar los datos de accidentalidad, con la accesibilidad territorial calculada sobre la malla vial, es indispensable ubicar geográficamente los eventos de accidentes que se tienen registrados, este procedimiento es lo que se conoce como georreferenciación, que es el proceso de asignarle un lugar a objetos geográficos, o que tienen dentro de sus atributos una posición particular, dentro de un sistema de coordenadas específico (Yao, 2009).

A partir de la información ordenada de la dirección en la base de datos, se utiliza la capacidad de OpenRefine de conectarse con varios servicios web para acceder al servicio del Google Maps API, al cual se le ingresa la dirección ya estandarizada y se almacenan las coordenadas que resultan de la búsqueda. Estas coordenadas están en el sistema de referencia geográfico WGS84, puesto que es uno de los sistemas más usados en el mundo, por ejemplo por los GPS.

Este procedimiento es particularmente útil hacerlo desde el OpenRefine, ya que permite procesar varias búsquedas en poco tiempo (algunas decenas por segundo), permitiendo la automatización de la tarea que se debe realizar para decenas de miles de registros de accidentes cuando se tienen varios años de información almacenada.

Posterior a la obtención y almacenamiento de las coordenadas, se procede a visualizar esta información utilizando un Sistema de Información Geográfica, como QGIS o ArcGIS, en el cual se introducen las coordenadas para otorgarle un lugar definido en el espacio a cada registro de evento almacenado. Con el

objetivo de manipular esta información en medidas de longitud y llevarlas a una escala cómoda para su edición, se transforma el sistema coordinado en el cual está representada y así visualizarla sobre la red vial de la ciudad.

De esta manera se obtiene un mapa que ilustra cada uno de los eventos registrados en la base de datos, con lo cual se ven todos los puntos esparcidos sobre la malla vial, hecho que no suele ser muy cómodo ni significativo visualmente para entender la información representada. Para superar este problema, se procede a generar una agregación de los datos de manera tal que se entienda su densidad en el espacio desde la visualización del mapa.

2.1.3 Agregación de datos y visualización

El cálculo de densidad de los accidentes tiene dos objetivos básicos: 1. Permitir una visualización significativa de la información, y 2. Interpolación, hasta donde se considere pertinente, la información sobre el territorio.

Para la agregación de los eventos registrados se usa el método de estimación de densidad de núcleo (Kernel Density Estimation: KDE), el cual permite producir mapas que son visualmente significativos a partir de puntos ubicados en el espacio con la información a mostrar (Anderson, 2009; Xie & Yan, 2008). Estos puntos pueden ser eventos en sí mismos o contener información adicional, como lo es el valor del NAE, que se puede utilizar para ponderar el impacto de unos eventos sobre otros.

La asignación de este valor de NAE a cada evento registrado es sencilla, puesto que solo depende de la gravedad registrada del accidente. De esta manera, solo se debe verificar que el atributo de 'gravedad' de la base de datos esté estandarizado y contenga únicamente los tres valores considerados: daños, heridos y muertos. Con esta verificación se procede a agregar automáticamente el atributo 'NAE' como un cálculo basado en otro atributo.

Uno de los parámetros principales de la estimación de densidad de núcleo es el radio de búsqueda de eventos, o ancho de banda de la función, que para el caso de los accidentes viales se determinó en 100 m. Esto responde a la longitud media de las cuadras de la ciudad, que en su centro histórico suele tener longitudes de 80 m, y medidas un poco mayores en zonas modernas de la ciudad. Al otorgarle esta medida al modelo de estimación, se declara que la influencia de un accidente puede darse hasta los 100 m.

Ahora bien, esta influencia no es completa y dependerá de la distancia a la que se encuentren los eventos registrados, como lo enuncia la primera ley de Tobler: “Todas las cosas están relacionadas entre sí, pero las cosas más próximas en el espacio tienen una relación mayor que las distantes” (Tobler, 1970).

Las funciones de probabilidad que determinan la influencia de un evento según su distancia, pueden ser muy variadas, entre las más usadas en el método de KDE están la uniforme, la triangular, la Epanechnikov, la cuadrática y la Gausiana (Jones, 1990). En este caso se usa la cuadrática, que tiene la forma mostrada en la Ecuación (2.1).

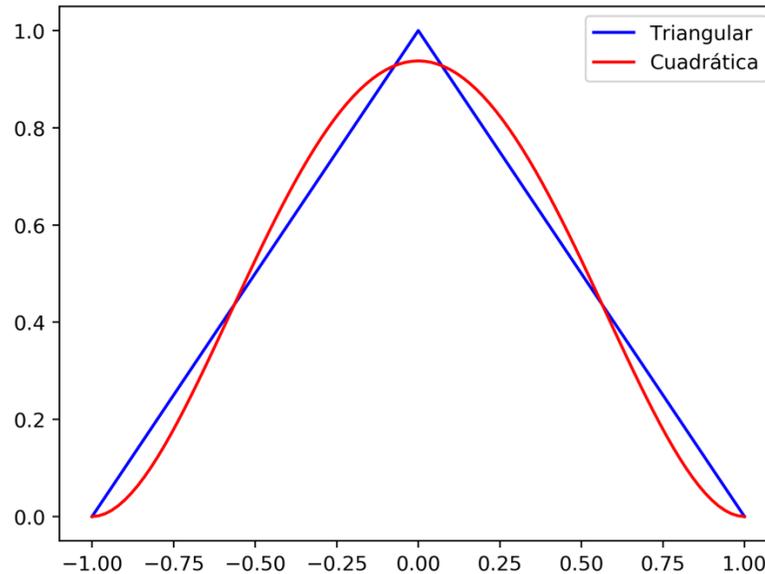
$$K(u) = \frac{15}{16}(1 - u^2)^2 \quad (2.1)$$

Se elige esta distribución porque es una que se reduce al llegar a los extremos, con lo que tiene en cuenta los eventos cercanos al límite del radio de búsqueda pero atenúa su influencia en la interpolación, y de otro lado, genera una reducción al 93% (15/16) del valor del evento donde se ubique y los dispersa en las cercanías de la ocurrencia. Este comportamiento se evidencia en la Gráfica 2-2 donde se compara la función de distribución cuadrática con la triangular, observándose el achatamiento en el centro y la dispersión contigua, terminando con una reducción del valor hacia los extremos.

De los dos componentes principales del KDE como método de suavización de información, el ancho de banda y la función de probabilidad, el que más afecta los

resultados es el ancho de banda, puesto que de este depende el nivel de agregación de los datos (Hastie et al., 2009).

Gráfica 2-2: Funciones de distribución de densidad.



Fuente: Elaboración propia

2.2 Accesibilidad

La manera en que se calcula la accesibilidad media global del territorio en este estudio es similar a lo que se ha hecho en otros trabajos para obtener esta medida sobre un territorio limitado a una ciudad (Escobar & García, 2012). El desarrollo del proceso comienza con la consecución de la red vial, la asignación de velocidades promedio a cada arco de la red, el cálculo de los tiempos promedio en cada nodo y la interpolación de estos tiempos para generar un mapa que cubra la totalidad del territorio.

2.2.1 Preparación de la red y asignación de velocidades

Para el primer paso, la red vial de una ciudad puede ser de fácil consecución, puesto que en la actualidad se cuenta con herramientas como OpenStreetMap que carga en internet esta información y se asigna un acceso público, así mismo,

los sistemas de información geográfica son tan populares y útiles actualmente, que incluso gobiernos de países en desarrollo como Colombia ya tienen publicados en sus páginas oficiales algunos recursos de este tipo.

El siguiente punto en el procedimiento es la asignación de velocidades promedio a cada arco de la red, estas se recogieron utilizando la metodología del vehículo flotante (De Fabritiis et al., 2008), que es simplemente la información en tiempo real de los vehículos monitoreados tomada con GPS, por lo que se define como un método móvil de captura de información, en contraste a los métodos fijos, que para registrar la información en la totalidad de la red son excesivamente ineficientes.

2.2.2 Cálculo de tiempos promedio

Al tener la red vial consolidada y las velocidades de operación asignadas a cada arco, se procede a calcular el tiempo que toma recorrer cada arco, dividiendo la longitud de este entre el valor de velocidad que tenga. Además, se debe verificar la direccionalidad de los arcos, debido a que este es un atributo real y funcional de las vías de la ciudad.

Posteriormente se procede a calcular la matriz de tiempos de viaje entre todos los nodos de la red. Se usa el algoritmo de caminos mínimos, tomando como variable a minimizar el tiempo calculado sobre cada arco. De esta manera se obtiene la matriz de tiempos que nos arroja principalmente dos resultados: los tiempos totales de cada nodo como origen y como destino.

En este trabajo se usan los tiempos de los nodos como destinos, a este vector se lo divide entre el número total de nodos menos uno para obtener el vector de tiempos promedio de viaje de todos los nodos de la red, que está acompañado de las coordenadas de cada uno, lo que permite ubicarlos en el espacio nuevamente para proceder a interpolar estos valores.

2.2.3 Interpolación

Teniendo los tiempos promedio calculados sobre cada nodo de la red, y a su vez estos ubicados en el espacio, se procede a aplicar el modelo de Kriging ordinario con semivariograma lineal para obtener la interpolación de los valores sobre el territorio. En este paso se obtiene un *raster* con la información interpolada, para el caso de este trabajo, el tamaño de celda de esta imagen pixelada es de 10m por 10m para que coincida con el tamaño del *raster* generado con la densidad de accidentes.

2.3 Correlación

Al tener los valores de accidentalidad y tiempos de accesibilidad esparcidos en el territorio, se procede a generar parejas con la ubicación geográfica como atributo en común. De esta manera se obtiene una lista de parejas de datos que servirá para calcular la correlación que puedan tener estas variables.

Con fines comparativos se calculan 3 coeficientes de correlación distintos: el de Spearman, el de Kendall y el de Pearson.

2.3.1 Spearman

El procedimiento para calcular este coeficiente consiste en ordenar los datos de cada variable y asignarles el valor de posición, con lo que se obtiene la misma lista de los elementos emparejados pero con valores ordinales. Al restar un valor de posición con su respectiva pareja, se obtiene la diferencia de posición o de rango y a partir de estos valores se aplica la ecuación (2.2) para obtener el valor del coeficiente de correlación de Spearman (r_s), donde (d_i) son las diferencias en los rangos y (n) es el tamaño de la muestra sobre la que se calcula (Siegel & Castellan, 1988).

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (2.2)$$

El cálculo del coeficiente de correlación de Spearman puede tener en cuenta la presencia de empates en el orden que se haga de los datos, otorgando a las observaciones empatadas el valor promedio de las posiciones que hubiesen obtenido con la asignación preliminar, y posteriormente aplicando la ecuación (2.3) para tener en cuenta estas alteraciones.

$$r_s = \frac{(n^3 - n) - 6 \sum_{i=1}^n d_i^2 - (T_X + T_Y)/2}{\sqrt{(n^3 - n - T_X)(n^3 - n - T_Y)}} \quad (2.3)$$

Siendo,

$$T_X = \sum_{j=1}^g t_j^3 - t_j \quad (2.4)$$

$$T_Y = \sum_{k=1}^h t_k^3 - t_k \quad (2.5)$$

Donde T_X y T_Y son los factores de corrección para las variables X y Y respectivamente; g y h son la cantidad de grupos empatados que exista para cada una de las variables; y t_j y t_k es el tamaño de cada grupo de elementos empatados.

2.3.2 Kendall

La manera en que se calcula el coeficiente de correlación de Kendall es similar al de Spearman, en cuanto a que se debe ordenar los valores de las variables y asignárseles los valores ordinales de posición. Posterior a esto se crea un conjunto de parejas ordenadas para cada variable, y sobre estos conjuntos se calcula la diferencia simétrica (Abdi, 2007).

$$\tau = 1 - \frac{2 \times [d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2)]}{n(n-1)} \quad (2.6)$$

De esta manera, se puede aplicar la ecuación (2.6) para obtener el valor del coeficiente de correlación de Kendall (τ), donde $(d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2))$ es la diferencia

simétrica de los conjuntos de parejas ordenadas (\mathcal{P}_1) y (\mathcal{P}_2) , y (n) es el tamaño de la muestra sobre la que se calcula o el número de observaciones total.

En las ocasiones en que la ordenación de los datos presenta empates y la posición de algunos elementos resulta la misma, el valor de posición de estos datos será el promedio de las posiciones que les hubiesen correspondido originalmente. Además, se debe aplicar la ecuación (2.7) para corregir el valor de este coeficiente.

$$\tau = \frac{n^2 - n - 2 \times [d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2)]}{\sqrt{(n^2 - n - T_X)(n^2 - n - T_Y)}} \quad (2.7)$$

Siendo,

$$T_X = \sum_{j=1}^g t_j^2 - t_j \quad (2.8)$$

$$T_Y = \sum_{k=1}^h t_k^2 - t_k \quad (2.9)$$

Donde T_X y T_Y son los factores de corrección para las variables X y Y respectivamente; g y h son la cantidad de grupos empatados que exista para cada una de las variables; y t_j y t_k es el tamaño de cada grupo de elementos empatados.

2.3.3 Pearson

Para calcular el coeficiente de correlación de Pearson se puede aplicar la ecuación (2.10)

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad (2.10)$$

Donde $(\rho_{X,Y})$ es el coeficiente de correlación de Pearson; (σ_{XY}) es la covarianza del conjunto de datos; (σ_X) es la desviación estándar de la variable X ; y (σ_Y) es la desviación estándar de la variable Y .

2.3.4 Coeficiente de determinación (R^2)

Al realizar una regresión lineal utilizando el método de los mínimos cuadrados, una medida que se obtiene de la correlación de los datos es el coeficiente de determinación, que se puede calcular con las ecuaciones mostradas a continuación.

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (2.11)$$

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \quad (2.12)$$

$$SS_{res} = \sum_i (y_i - f_i)^2 \quad (2.13)$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (2.14)$$

Donde \bar{y} es el promedio del conjunto de datos observados; y_i es cada dato; n es el número total de datos; SS_{tot} es la suma de los cuadrados de la diferencia entre cada registro y el promedio de los datos; SS_{res} es la suma de los cuadrados de los valores residuales, obtenidos de la diferencia entre cada observación y su valor según la función obtenida de la regresión; y R^2 es el coeficiente de determinación.

Si ya se ha calculado el coeficiente de correlación de Pearson, el coeficiente de determinación se puede computar como el cuadrado de este.

3. Resultados y análisis

En este capítulo se mostrarán los resultados obtenidos al aplicar los conceptos explicados con anterioridad a un caso particular de estudio. En este caso se usará la ciudad de Manizales, lugar donde se ubica la sede de la universidad desde la que se produce este proyecto. En este sentido, se describirá la ciudad de Manizales de manera general y se hará hincapié en sus condiciones de movilidad, puesto que son pertinentes para entender las particularidades de los datos analizados.

Manizales es un municipio de Colombia que se ubica en el departamento de Caldas, en la región andina del país, que es una zona montañosa, otorgándole una topografía de pendientes fuertes al trazado vial de la ciudad. El municipio cuenta con un área dentro del perímetro urbano de 59,68 km² y una población de 371.345 habitantes para 2017 (DANE, 2016), de este modo dando una densidad urbana poblacional media de 6.222 hab./ km².

Desde el punto de vista del desarrollo urbanístico, la ciudad presenta una dinámica bicéntrica en las actividades financieras y comerciales principales, consecuencia natural de ser una ciudad mediana, tendencia que tienen la gran mayoría de aglomeraciones en una etapa transicional de su crecimiento (Louf & Barthelemy, 2013). Este aspecto hace que su desarrollo vial sea alargado, en este caso, en sentido occidente-oriente.

Dentro de las falencias que tiene la ciudad en el apartado de movilidad está el crecimiento acelerado del parque automotor, la alta accidentalidad, la discontinuidad en las mejoras de la infraestructura y la falta de políticas definidas e incluyentes de movilidad en general. Estas deficiencias en la gestión de la

movilidad municipal se ven reflejadas en sus índices de accidentalidad, que si bien están por debajo de la media colombiana, son muy altos en el contexto mundial.

3.1 Accidentalidad

La base de datos que se obtiene de la secretaría de tránsito tiene consolidada la información desde el año 2007 hasta agosto del 2016 inclusive, y presenta datos para cada registro de accidente como la fecha, la clase de accidente, la dirección, si se presenta en la zona urbana o en la periferia rural de la ciudad, si el sector es comercial o residencial y la gravedad de las consecuencias, donde se enuncia si hubo víctimas fatales, heridos o solo daños materiales. Es de esta última característica de donde se obtiene el valor de NAE de cada evento para luego ser agregado.

En la Tabla 3-1 se muestra la cantidad de accidentes registrados en la base de datos separados por la gravedad del evento, en esta se observa que la gran mayoría de accidentes ocasionan daños materiales, ya sea entre los vehículos involucrados o al mobiliario urbano circundante. La ocurrencia de heridos, aunque menor, tiene una representación importante en el total de eventos registrados, contando con cerca del 30% de la repartición. Y por último, la proporción de accidentes que involucran muertes con respecto al total de accidentes puede verse como bastante baja, pero debido al costo social que esto representa (Elvik, 2000), y teniendo en cuenta que en un evento puede haber varias víctimas fatales, este resultado no deja de ser una consecuencia importante de la dinámica de movilidad en la ciudad.

Esta base de datos se limpió con la ayuda de estudiantes del semillero de investigación de la oficina de movilidad de la Universidad Nacional de Colombia Sede Manizales, que se dedicaron a homogenizar la información de la dirección de cada registro para que el procesamiento de geocodificación tuviera éxito. Al tener esta información estandarizada se procede a pasar la base de datos por el

software OpenRefine, que se encarga de buscar las coordenadas geográficas de los registros basado en la dirección registrada.

Tabla 3-1: Distribución de accidentes por gravedad.

Gravedad	Cantidad	Proporción
Daños	16704	69.04%
Heridos	7415	30.65%
Muertos	76	0.31%
Total	24195	100.00%

Fuente: Elaboración propia con datos suministrados por la secretaría de tránsito

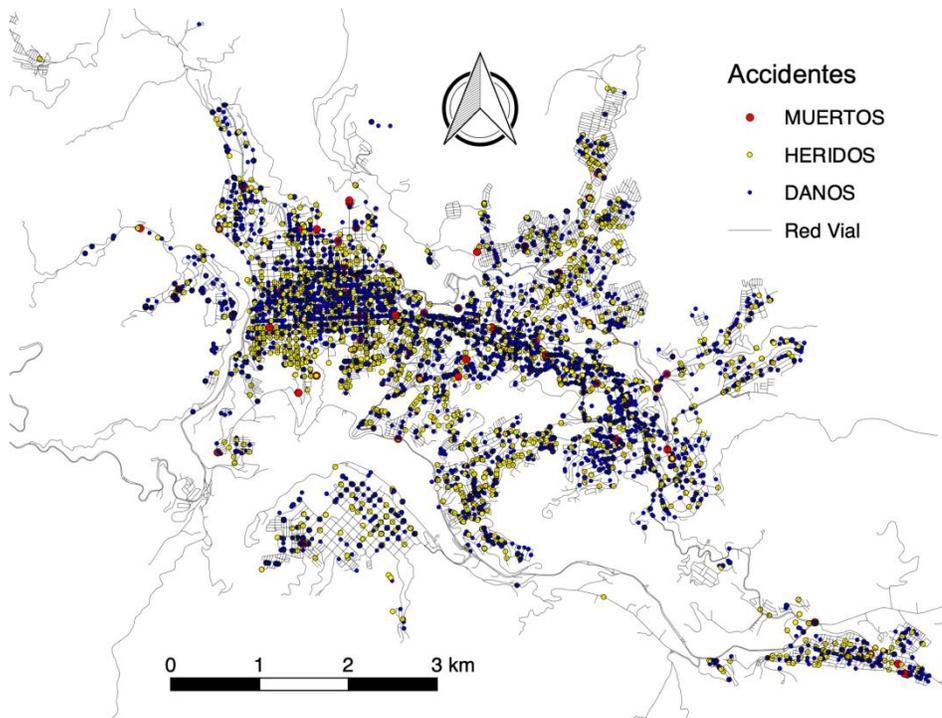
Este procedimiento tuvo éxito en el 97% de los casos, obligando a ubicar manualmente el resto de los incidentes. Este pequeño obstáculo da pie para proponer una mejor metodología de registro de accidentes, una en la cual las autoridades de tránsito tengan a acceso a herramientas que les permitan guardar la ubicación geográfica al momento de hacer el levantamiento del siniestro. Esto se puede lograr fácilmente con la utilización de unas tablets o dispositivos móviles con GPS integrado y el desarrollo de una aplicación tipo formulario. De esta manera se incurriría en la optimización del registro de la información, y su costo de implementación es muy bajo haciendo el balance con la información acertada y pertinente que otorga a los analistas de estas bases de datos.

Teniendo todos los registros de la base de datos con sus respectivas coordenadas geográficas, se puede proceder a ubicarlos espacialmente en un mapa, superpuestos con la red vial de la ciudad, para obtener una visualización y una noción de de su distribución sobre el territorio municipal.

En la Figura 3-1 se ilustran los accidentes registrados en la ciudad de Manizales entre los años 2007 a 2016, que servirán de insumo para este caso de estudio. En esta figura se puede observar de manera preliminar la densidad que existe de estos eventos hacia el centro de la ciudad y sobre las avenidas principales. Observaciones como esta permiten pensar en la posibilidad de relación entre la accesibilidad y la accidentalidad que se pretende evaluar.

A pesar de que en términos generales se nota una mayor densidad hacia el centro y las troncales de la ciudad, si se observan los puntos rojos que representan los eventos con víctimas fatales, se puede percibir un mayor esparcimiento, lo que posiblemente redunde en una menor relación con la accesibilidad territorial por parte de estos registros.

Figura 3-1: Accidentes en la ciudad de Manizales, 2007-2016

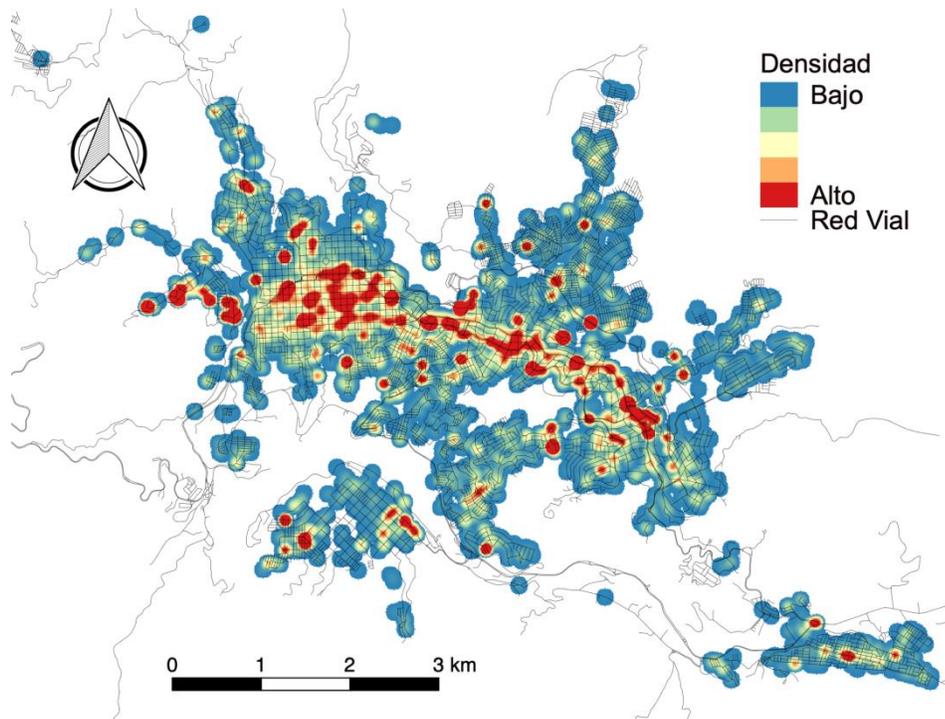


Fuente: Elaboración propia

A partir de estos registros ubicados espacialmente, se procede a calcular sobre el territorio un mapa de densidades basándose en la estimación de densidad de núcleo (Kernel Density Estimation: KDE), y utilizando como función de distribución la función cuadrática mencionada en la ecuación (2.1). Los mapas obtenidos, llamados “Raster”, se pueden entender como imágenes pixeladas, donde cada unidad o píxel contiene la información calculada. En este caso se usaron tamaños de píxel de 10m de alto y ancho, que corresponden al 10% del radio de búsqueda del KDE.

En la Figura 3-2 se observa el mapa de calor resultante de calcular la estimación de densidad de núcleo (Kernel Density Estimation), teniendo en cuenta cada evento de accidente como la unidad. En este mapa se puede ver que la mayor concentración de accidentes se da hacia el centro histórico de la ciudad y sobre las avenidas que lo conectan con el centro de negocios moderno.

Figura 3-2: Mapa de calor de accidentes simples



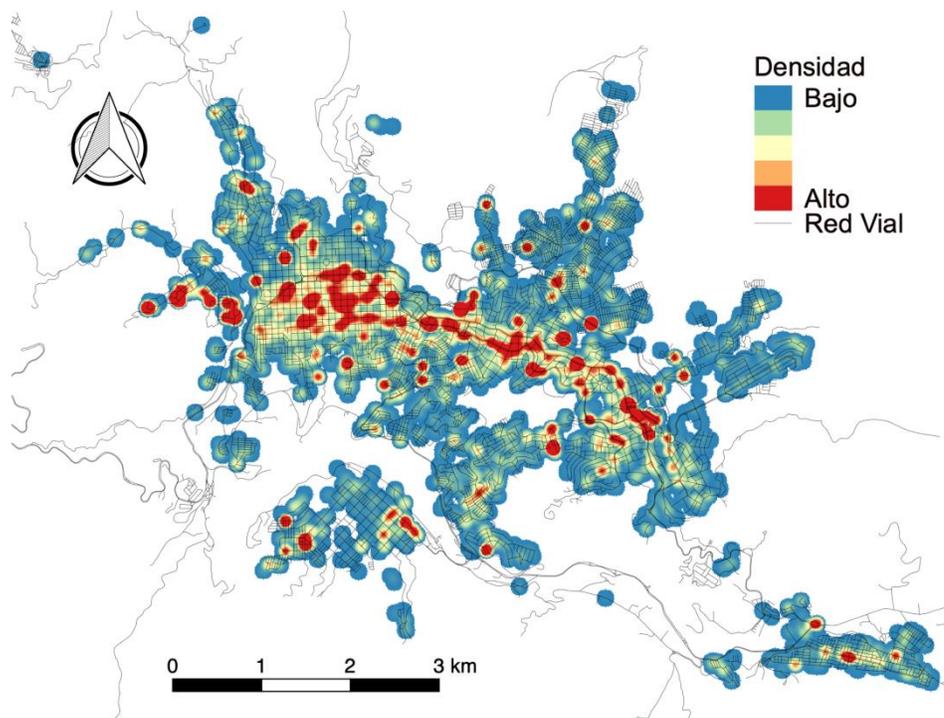
Fuente: Elaboración propia

En contraste, la Figura 3-3 muestra el resultado del mismo proceso pero otorgándole el valor de NAE a cada evento, de manera tal que los eventos de solo daño representan la unidad, los registros de heridos tienen un peso de 1.43 y por último, aquellos que involucraron víctimas fatales, tienen un peso de 7.48. Estos valores responden a la eq tal del NAE.

De esta manera, al comparar ambos mapas se podría observar el impacto que tienen los sectores con ocurrencia de muerte, pero en este caso, los cambios son muy sutiles. Esto se debe a la alta cantidad de los eventos con daños y a la vez a

los pocos eventos con víctimas fatales presentes en la base de datos, con lo que el impacto de estos últimos se disuelve con la sumatoria de los primeros.

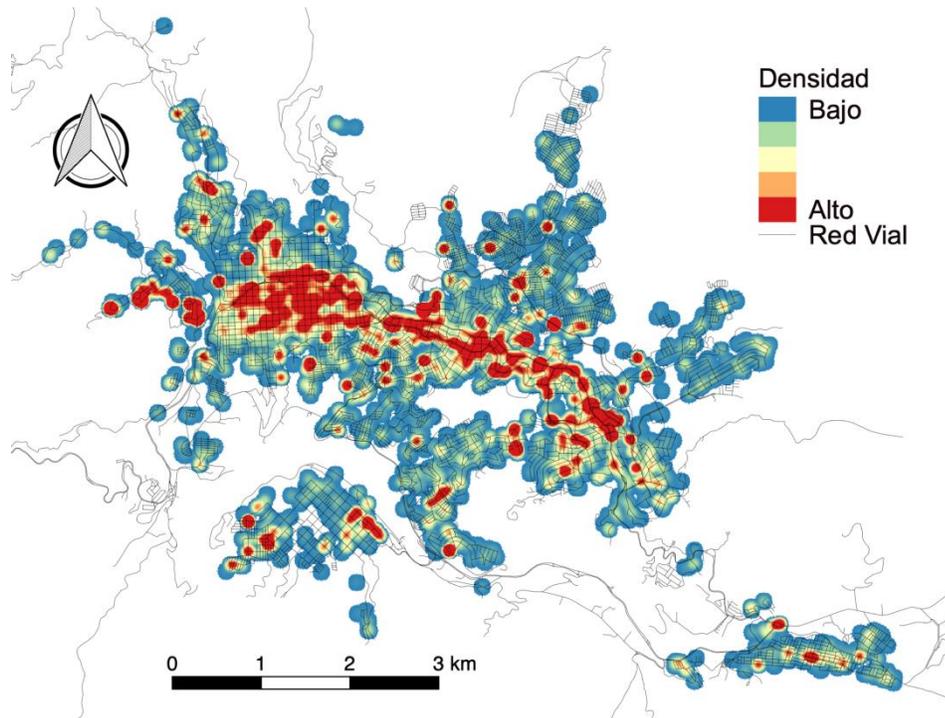
Figura 3-3: Mapa de calor de accidentes con NAE



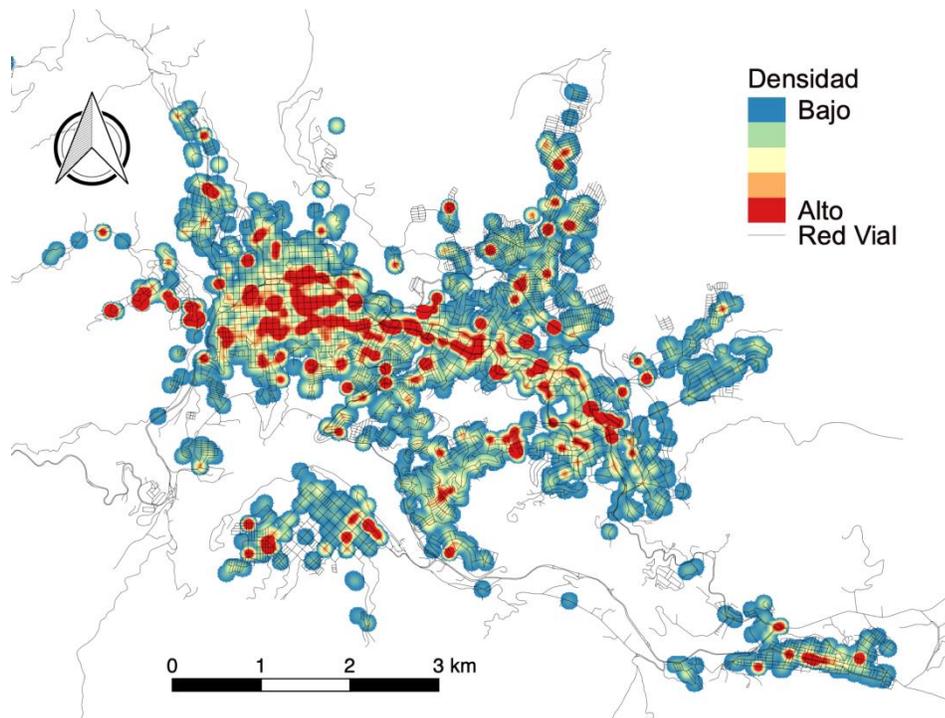
Fuente: Elaboración propia

Así las cosas, se hace pertinente obtener valores de cada tipo de gravedad de accidente para comparar con estos valores combinados y verificar la posibilidad de que unos se relacionen más directamente con la accesibilidad que otros. Bajo esta premisa se procede a calcular los mapas de densidad usando los datos de los accidentes con daños, heridos y muertes por separado.

En la Figura 3-4 se ilustra el mapa de densidad de los registros con solo daños, en este se observa una mayor homogenización de la densidad en la zona centro de la ciudad. Por otro lado en la Figura 3-5 se muestra el mapa de densidad usando los eventos con heridos de la base de datos, en este se observa una similitud mayor con los mapas anteriores que usan las tres clases de gravedad.

Figura 3-4: Densidad de accidentes con daños

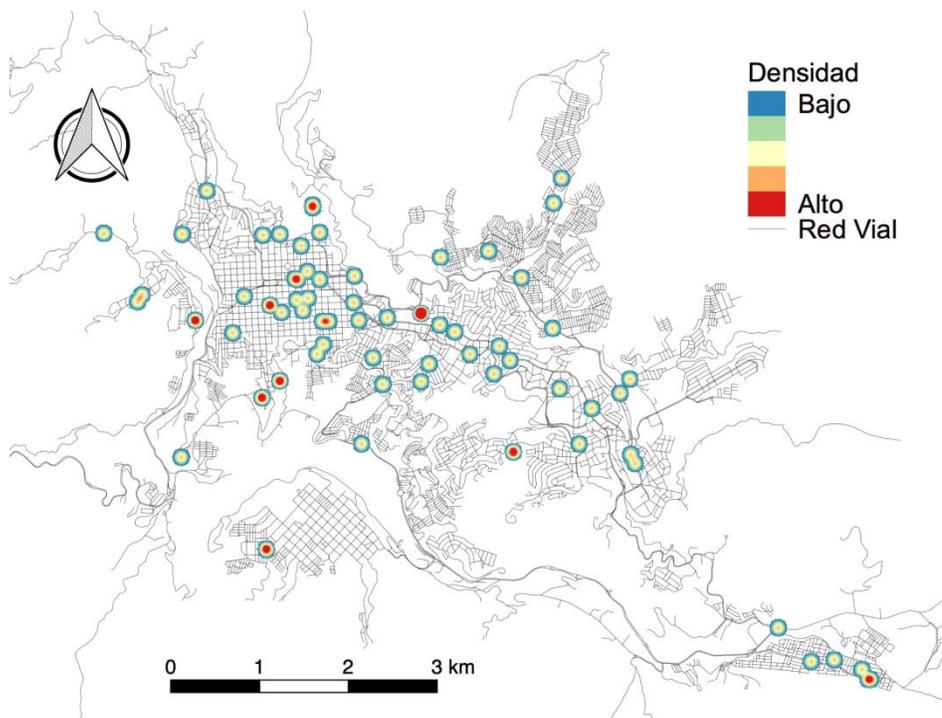
Fuente: Elaboración propia

Figura 3-5: Densidad de accidentes con heridos

Fuente: Elaboración propia

En la Figura 3-6 que presenta la densidad de accidentes con víctimas fatales lo primero que se percibe es la poca cantidad de eventos que hay en la ciudad. En algunos de los puntos se evidencia la existencia de dos o más eventos, estos lugares, solo con este insumo informativo, deberían ser de alta prioridad para la administración municipal.

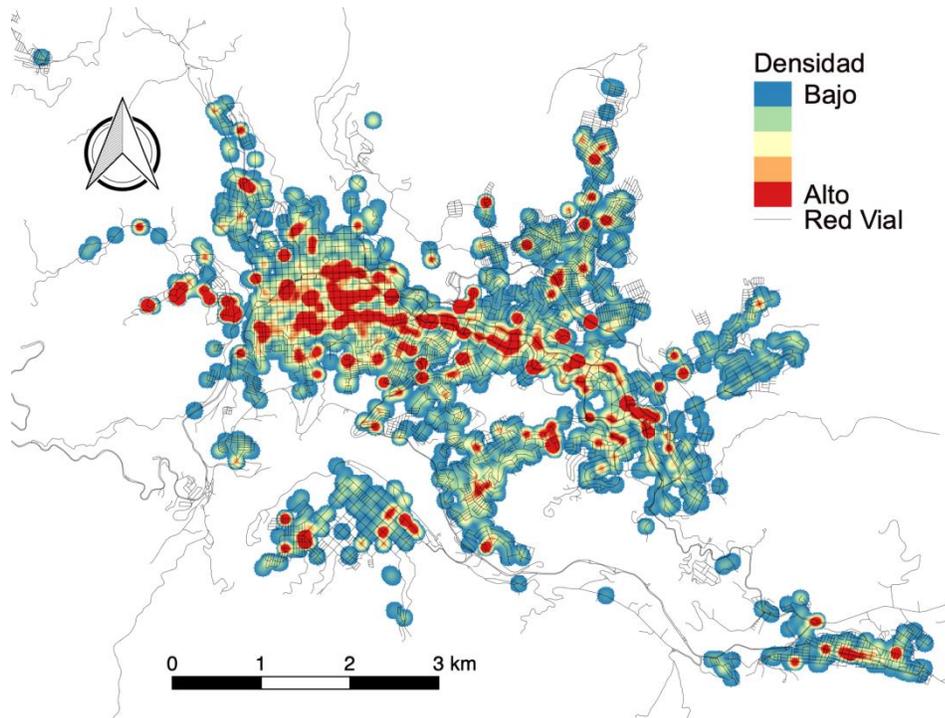
Figura 3-6: Densidad de accidentes con víctimas fatales



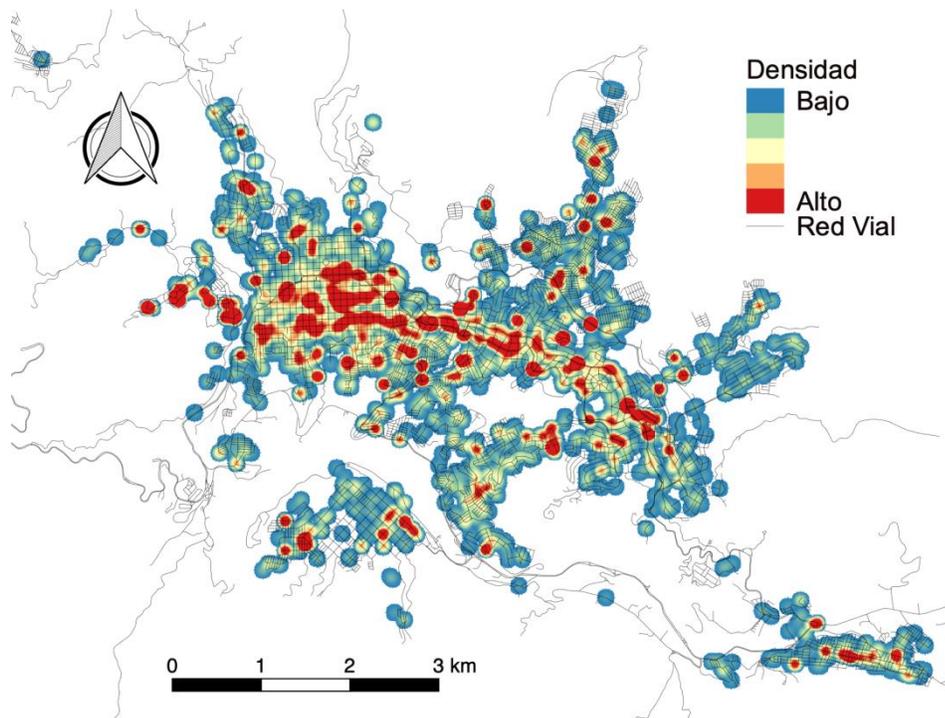
Fuente: Elaboración propia

Con el ánimo de prescindir de estos casos fatales de accidentes, se procede a calcular también mapas de densidad usando solamente los registros de heridos y muertos, tanto en valores simples como teniendo en cuenta el valor NAE.

El mapa de densidad con los valores simples se puede observar en la Figura 3-7, y el mapa afectado con el NAE de 1.43 para los casos con heridos se ilustra en la Figura 3-8. Ambos mapas son muy similares, salvo unos pocos focos donde los heridos tienen una presencia mayor cuando son afectados por el NAE.

Figura 3-7: Densidad heridos y muertos

Fuente: Elaboración propia

Figura 3-8: Densidad heridos y muertos con NAE

Fuente: Elaboración propia

En todos los casos de los mapas es evidente que la ocurrencia de accidentes se vuelven más dispersa hacia la periferia, particularmente en las vías perimetrales con ausencia de intersecciones.

Cada uno de estos mapas se relacionará con los datos de accesibilidad calculados a continuación y se obtendrá una medida de cada uno de esos apareamientos para evaluar su correlación.

3.2 Accesibilidad

Con el objetivo de relacionar los datos de accidentalidad con los tiempos medios de viaje de la red vial de la ciudad, se procede a calcular estos tiempos y llevarlos a un formato que sea compatible con los mapas de densidad anteriormente calculados, esto es, a un *raster* con celdas de 10m por 10m.

El primer paso de este procedimiento es tener a punto la red vial a utilizar, lo que comprende verificar la continuidad y tener asignada la velocidad promedio de operación de cada arco. Este paso se cumple con facilidad en el presente trabajo pues la red vial se mantiene actualizada en la oficina de movilidad de la universidad, lo que permite hacer una verificación rápida de estas condiciones y tenerla lista para el cálculo posterior.

Luego de tener la red lista, se procede a calcular el tiempo mínimo que toma ir de un nodo a otro de la red, para todos los nodos que la componen. Esto nos arroja la matriz de tiempos con todos los nodos de la red, de la cual se toma el promedio de las columnas para obtener el tiempo promedio que toma llegar a cada uno de los nodos.

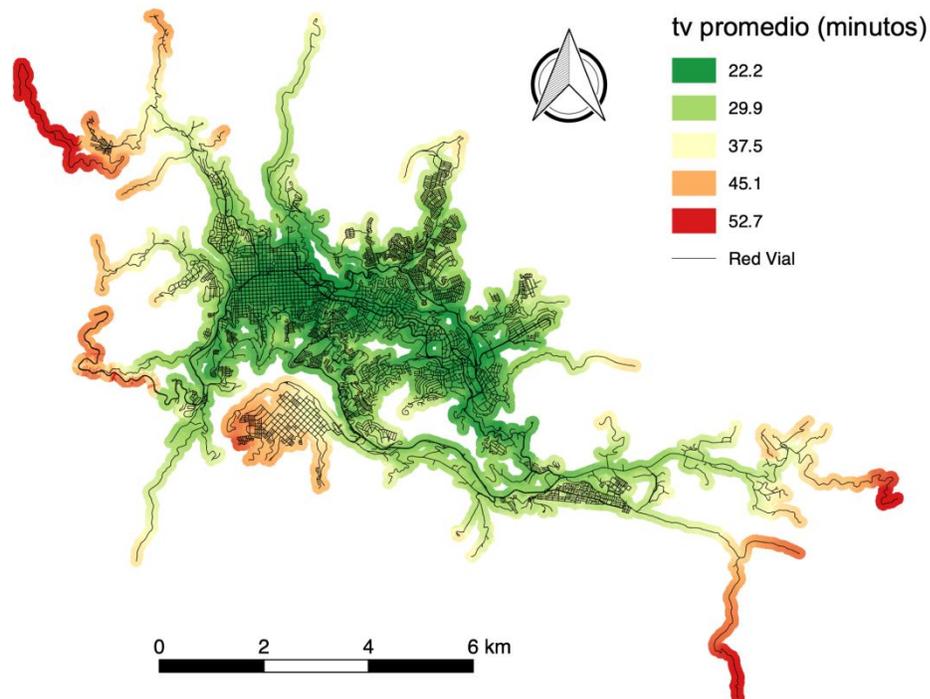
Estos tiempos, pertenecientes a puntos con una ubicación espacial se despliegan sobre la red vial. Posteriormente estos datos se interpolan usando el método de Kriging ordinario y un semivariograma lineal.

En la Figura 3-9 se observa la interpolación de los datos de tiempo calculados sobre la red vial de la ciudad. En esta se evidencia el rápido acceso que suele

tener el centro de la ciudad, con tiempos promedio inferiores a los 20 minutos, zonas donde también se perciben mayores las densidades de accidentalidad.

De otro lado, las vías de la periferia donde los tiempos promedio de viaje se aproximan a los 60 minutos, son las zonas donde se registraron solo unos pocos accidentes.

Figura 3-9: Raster con tiempos de viaje promedio



Fuente: Elaboración propia

3.3 Correlación

Para el análisis combinado de las dos variables calculadas hasta ahora se procede a crear parejas de datos con el tiempo medio de viaje de la accesibilidad y con cada caso de la densidad de accidentalidad. Para cada uno de los casos mostrados a continuación se ilustra la forma de la regresión lineal por mínimos cuadrados usando el logaritmo del tiempo, que se ajusta un poco mejor que la función lineal.

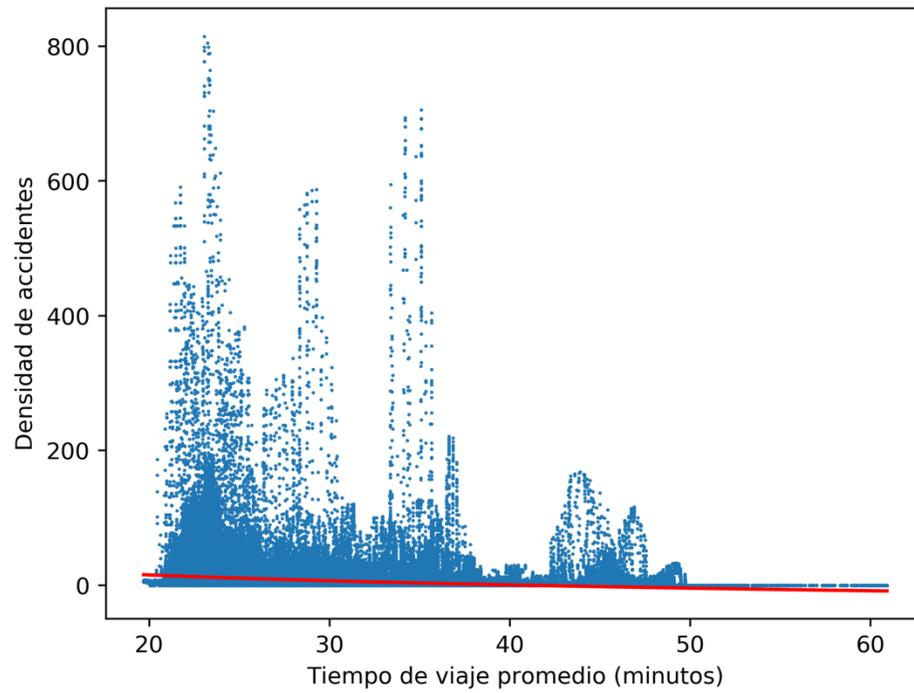
En la Gráfica 3-1 se observan los datos de la densidad de accidentes, sin afectar con el NAE, contra la información de tiempo, evidenciando que en los menores tiempos de accesibilidad de la red se encuentran densidades más altas que en las zonas con mayores tiempos. Sin embargo, son tantos los datos que se aproximan a 0 que los pocos valores altos no rigen el comportamiento de la regresión.

Las ecuaciones (3.1) y (3.2) muestran las funciones de la regresión lineal, usando el tiempo y el logaritmo del tiempo respectivamente.

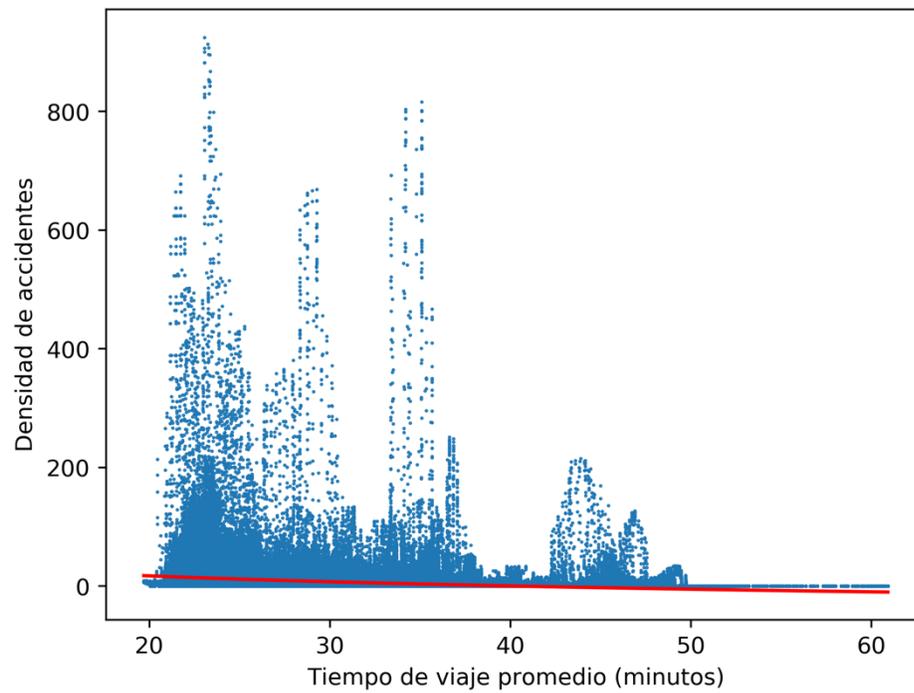
$$DA = -0.553 * t + 23.28 \quad (3.1)$$

$$DA = -21.37 * \ln(t) + 79.15 \quad (3.2)$$

En la Gráfica 3-2 se muestra la dispersión de datos de la densidad de accidentes, afectada por el valor del NAE, emparejados con su respectivo tiempo sobre el territorio. En esta se puede observar un comportamiento similar a la gráfica anterior, cambiando solamente en la magnitud de la densidad debido al escalamiento por los factores de NAE usados.

Gráfica 3-1: Densidad accidentes simples vs tiempo de viaje promedio

Fuente: Elaboración propia

Gráfica 3-2: Densidad de NAE vs tiempo de viaje promedio

Fuente: Elaboración propia

Las ecuaciones (3.3) y (3.4) muestran las funciones de la regresión lineal, usando el tiempo y el logaritmo del tiempo respectivamente, con los datos de accidentes afectados por el NAE.

$$DA = -0.634 * t + 26.74 \quad (3.3)$$

$$DA = -24.51 * \ln(t) + 90.82 \quad (3.4)$$

De otro lado, al realizar un análisis similar a los datos de accidentes en cada categoría de gravedad, se empiezan a evidenciar algunas diferencias en la distribución de las dispersiones.

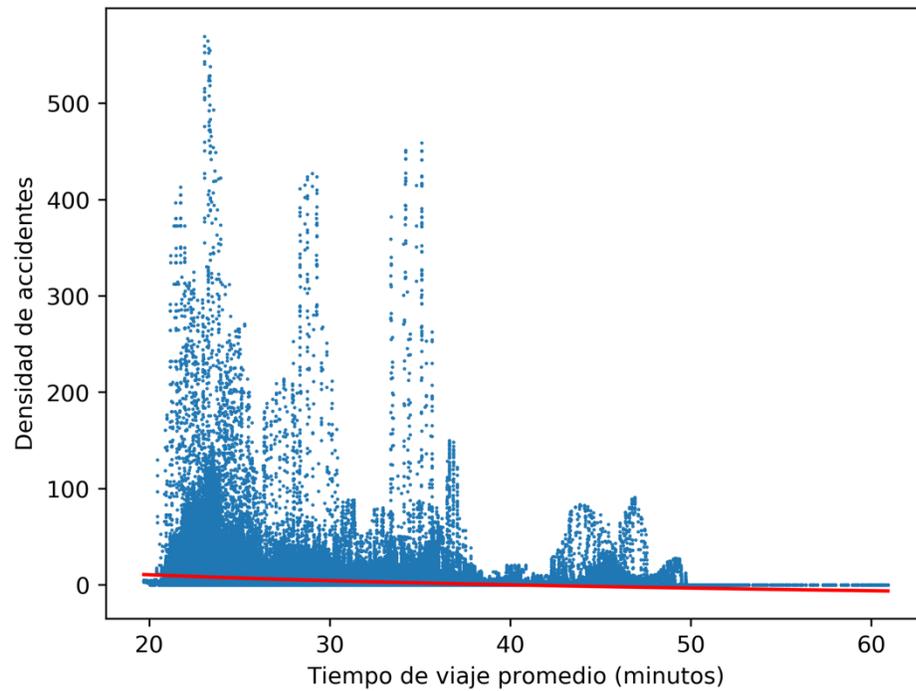
Para los datos de accidentes de solo daños mostrados en la Gráfica 3-3, se puede evidenciar, además de la reducción en magnitud de la densidad, un cambio considerable en la altura relativa del cúmulo alrededor de los 44 minutos de tiempo, reduciéndose con respecto a las gráficas anteriores. De esto se puede deducir que alrededor de estos tiempos deben existir un número importante de accidentes con heridos y/o muertos que afectan en el alza al valor combinado por encima de otros grupos de tiempo.

En las ecuaciones (3.5) y (3.6) se observan las regresiones por mínimos cuadrados de este conjunto de datos.

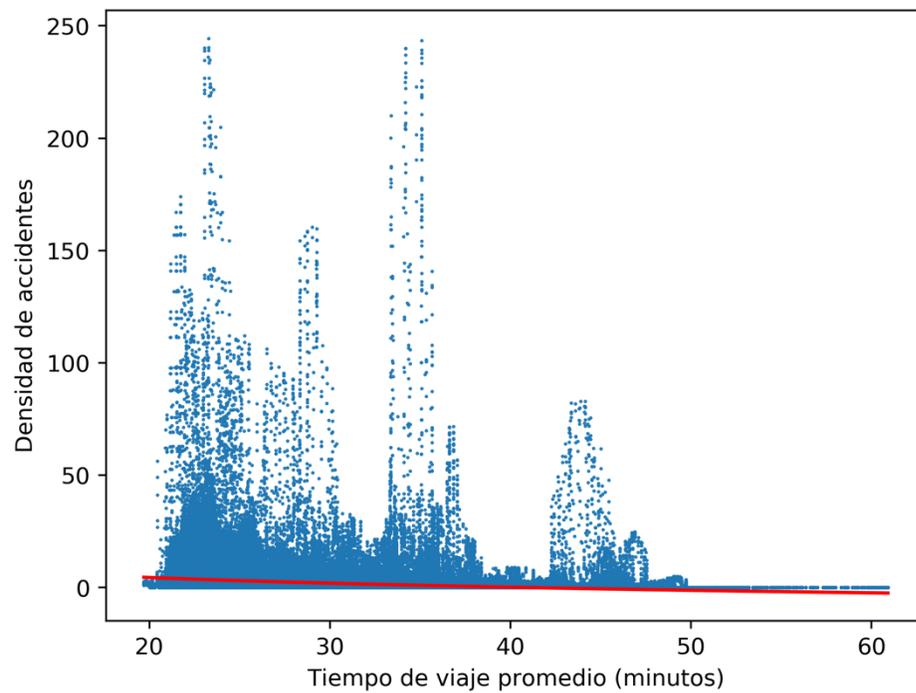
$$DA = -0.389 * t + 16.33 \quad (3.5)$$

$$DA = -15.09 * \ln(t) + 55.80 \quad (3.6)$$

En cuanto a la dispersión de los accidentes con heridos ilustrada en la Gráfica 3-4, se puede evidenciar también un comportamiento similar al de las dispersiones con las tres categorías salvo alrededor del mismo rango de tiempos de la alteración de la gráfica anterior, pero esta vez los valores están al alza, lo que confirma la sospecha de que en esta zona, al tener valores bajos de accidentes con solo daños, debían ser altos, de forma relativa, los valores de heridos y muertos.

Gráfica 3-3: Densidad accidentes con daños vs tiempo de viaje promedio

Fuente: Elaboración propia

Gráfica 3-4: Densidad accidentes con heridos vs tiempo de viaje promedio

Fuente: Elaboración propia

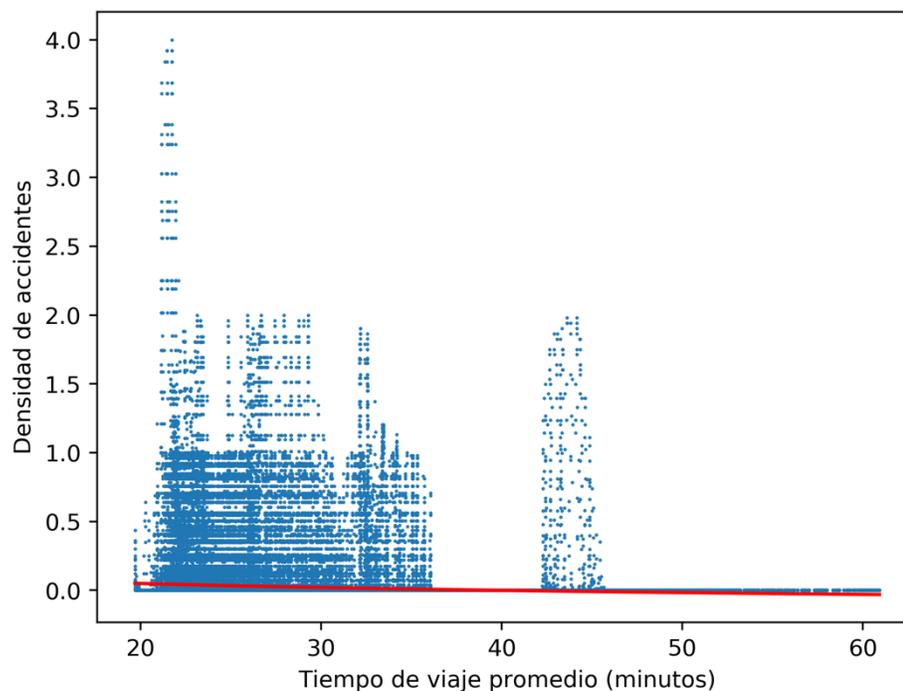
Las fórmulas de las regresiones lineales usando el tiempo y el logaritmo del tiempo se muestran en las ecuaciones (3.7) y (3.8) respectivamente, para los datos de los accidentes con heridos.

$$DA = -0.161 * t + 6.88 \quad (3.7)$$

$$DA = -6.21 * \ln(t) + 23.10 \quad (3.8)$$

En la Gráfica 3-5 que ilustra la dispersión de accidentes con víctimas fatales se evidencia un comportamiento muy distinto al resto, principalmente debido a que esta cuenta con muchos menos registros. En esta se puede percibir que los picos de los cúmulos llegan a valores de 4, 2 y 1.

Gráfica 3-5: Densidad accidentes con víctimas fatales vs tiempo de viaje promedio



Fuente: Elaboración propia

De todas las ecuaciones de regresión, las ecuaciones (3.9) y (3.10) que pertenecen a la regresión usando el tiempo y logaritmo del tiempo cruzando con los valores de densidad de accidentes fatales, muestran los valores más bajos de

pendiente debido a que esta es la dispersión que presenta la mayor cantidad de valores en cero a todo lo largo del rango tiempo.

$$DA = -0.002 * t + 0.08 \quad (3.9)$$

$$DA = -0.07 * \ln(t) + 0.27 \quad (3.10)$$

Por último, se analizan las dispersiones de los eventos con heridos y con muertes combinados, tanto en términos simples como ponderados por el valor del NAE. La primera se ilustra en la Gráfica 3-6, y la segunda en la Gráfica 3-7, presentando ambas un comportamiento en su distribución muy parecido, cambiando básicamente en la magnitud de sus valores al haber afectado los eventos con heridos por el factor de 1.43.

En las ecuaciones (3.11) y (3.12) se muestran las regresiones usando el tiempo y el logarito del tiempo respectivamente con relación a los accidentes de heridos y muertes sin ponderación del NAE.

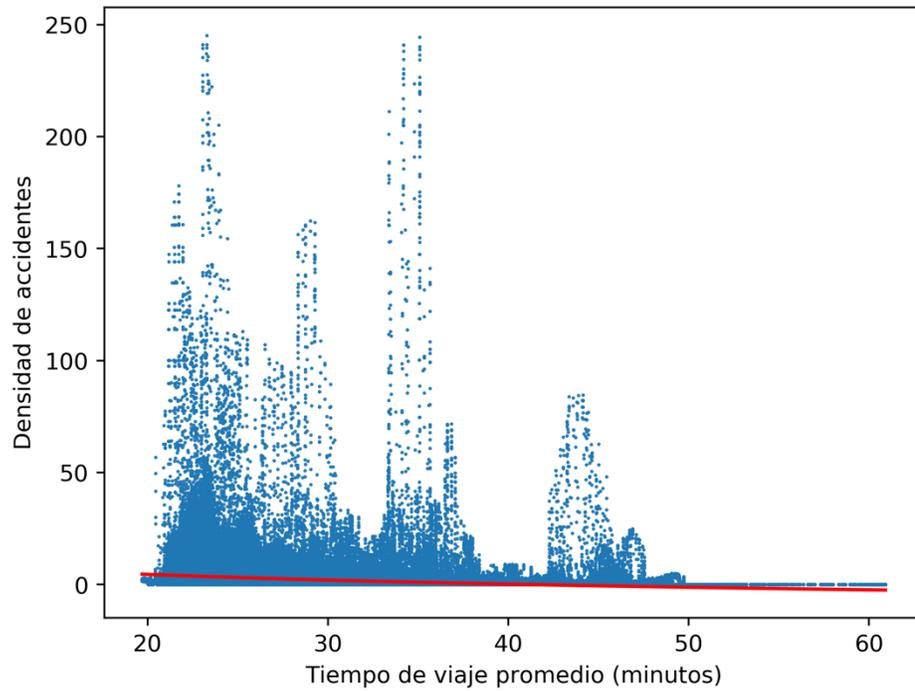
$$DA = -0.163 * t + 6.95 \quad (3.11)$$

$$DA = -6.29 * \ln(t) + 23.36 \quad (3.12)$$

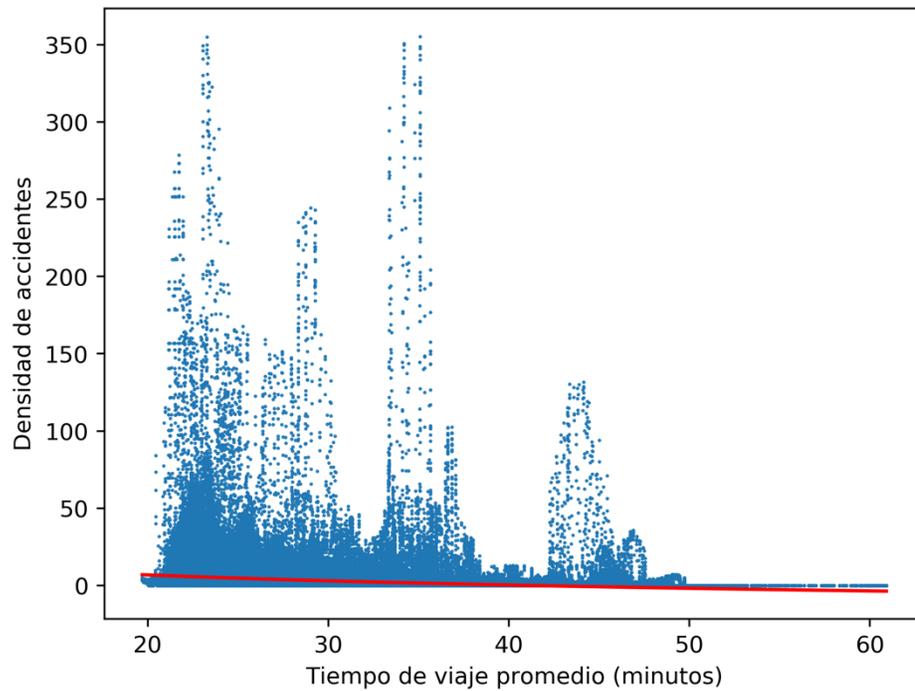
En las ecuaciones (3.13) y (3.14) se muestran las funciones obtenidas de la regresión lineal usando el tiempo y el logaritmo del tiempo respectivamente para los valores de densidad de accidentes de heridos y muertos afectados por el NAE.

$$DA = -0.245 * t + 10.41 \quad (3.13)$$

$$DA = -9.43 * \ln(t) + 35.03 \quad (3.14)$$

Gráfica 3-6: DA con muertes y heridos vs tiempo de viaje promedio

Fuente: Elaboración propia

Gráfica 3-7: DA con muertes y heridos con NAE vs tiempo de viaje promedio

Fuente: Elaboración propia

Apartir de estos distintos conjuntos de datos analizados, se procede a calcular los coeficientes de correlación pertinentes para todos los casos. En la tabla Tabla 3-2 se muestran los resultados obtenidos para cada conjunto de datos y usando los distintos métodos.

Tabla 3-2: Coeficientes de correlación con distintos métodos y conjuntos de datos

Método	Spearman	Kendall	Pearson		Coef. Deter.	
			DA vs t	DA vs log(t)	DA vs t	DA vs log(t)
DA						
Simples	-0.5449	-0.4178	-0.1857	-0.2081	0.0345	0.0433
NAE	-0.5452	-0.4179	-0.1858	-0.2082	0.0345	0.0433
Daños	-0.5366	-0.4124	-0.1875	-0.2105	0.0352	0.0443
Heridos	-0.5130	-0.3949	-0.1765	-0.1969	0.0312	0.0388
Muertes	-0.2123	-0.1718	-0.1238	-0.1390	0.0153	0.0193
MyH Simple	-0.5151	-0.3963	-0.1771	-0.1976	0.0314	0.0390
MyH NAE	-0.5157	-0.3968	-0.1791	-0.1998	0.0321	0.0399

Fuente: Elaboración propia

En primer lugar, se muestran los coeficientes de correlación por rangos de Spearman, que fueron los más altos de entre todos los calculados, con valores para las observaciones combinadas de entre -0.51 y -0.54, lo que se puede interpretar como una relación inversa y suficientemente alta para considerarse pertinente. La dirección de la relación es predecible en cuanto que a menores tiempos de accesibilidad se esperaban mayores valores de densidad de accidentes.

Inicialmente era de esperarse que tanto los coeficientes de Spearman y Kendall fueran superiores a los coeficientes de Pearson debido a que estos describen solo la tendencia de los datos a moverse en un sentido de manera conjunta, mientras que el de Pearson, además de esto, es más riguroso al reflejar el comportamiento con respecto a una función particular, lo que deriva usualmente en valores inferiores.

Debido a este mismo principio, se tiene que el coeficiente de correlación de Pearson puede variar según se afecte las variables observadas, por ejemplo al

calcularse con respecto al logaritmo del tiempo, ajustándose mejor o peor a la función que lo rija. De otro lado, como tanto el coeficiente de Spearman y el de Kendall transforman los valores de las variables a sus posiciones ordenadas, son independientes de las transformaciones que se les pudieren realizar.

Al compararse los coeficientes de Spearman y Kendall, se observa que los primeros son mayores. Esto permite interpretar que los conjuntos de datos, en términos generales, no presentan valores atípicos extremos, puesto que si este fuese el caso, los coeficientes de Spearman serían menores debido a que este es más sensible a la presencia de estas particularidades y resultaría en un menor valor que el de Kendall.

Otra particularidad de estos resultados son los valores de los coeficientes calculados a los accidentes con muertes, que resultaron en valores muy inferiores a todos los demás. Esto se debe principalmente al número de registros que se tenían en esta categoría de accidentes, puesto que muchas parejas de observaciones resultaron con valores de 0 en la densidad de accidentes, con lo que las variaciones en el tiempo promedio de la accesibilidad no se acompañaban con ningún tipo de variación en la accidentalidad.

4. Conclusiones y recomendaciones

4.1 Conclusiones

Las herramientas de manipulación de bases de datos son claves para lograr el procesamiento eficiente de grandes conjuntos de información que deban ser analizados. El percance que se tiene actualmente, es que las conexiones en internet, para el procesamiento o adquisición de información desde un servicio web, pueden llegar a ser lentas cuando se escala a un volumen considerablemente grande de solicitudes.

Posiblemente en el futuro cercano se superen estos obstáculos con la implementación de las velocidades 5G que ya empiezan a surgir en el mercado. Esto permitirá que se abaraten los procesos y hará que servicios como el de Google Maps pueda expandir los límites de búsquedas que tiene actualmente.

De los resultados obtenidos se puede decir que uno de los parámetros que más los afecta es el ancho de banda o radio de búsqueda de la función de distribución usada en la estimación de densidad de núcleo. Escogiendo un valor mayor para esta distancia, los valores de densidad de los accidentes pueden ser más suavizados, más dispersos, lo que posiblemente repercutiría en una mejor correlación con los tiempos de accidentalidad. Queda para una próxima investigación variar estos valores y verificar esta hipótesis, teniendo cuidado de mantener el objetivo y la pertinencia del resultado.

Desde el punto de vista del cumplimiento del objetivo del trabajo, que es identificar zonas de la ciudad con alto índice de accidentalidad a partir del análisis

de accesibilidad de la red, se puede decir que lo que se encontró es que esto no se puede lograr debido a las diferentes escalas con que se analiza cada variable.

Es decir, tratar de encontrar intersecciones puntuales con alta accidentalidad a partir de la información de tiempos de la red, es imposible. Lo más cerca que se puede hacer es inferir en los sectores de la ciudad que pueden tener una alta accidentalidad debido a la alta accesibilidad que presentan. De nuevo, esto se debe a la escala con que se interpolan los datos, ya que la información de tiempo es mucho más suavizada que la densidad de accidentes.

Así mismo, a pesar de que la correlación entre las variables analizadas es moderadamente buena, pueden existir otras que se relacionen mejor con la accidentalidad, especialmente factores que se ajusten a una incidencia cercana, como lo puede ser el flujo vehicular de la ciudad o características físicas del entorno que generan accidentes. Estos son factores que pueden ser menos homogéneos en el territorio y por tanto pueden tener unas variaciones tan fuertes como la accidentalidad evaluada.

4.2 Recomendaciones

Con la experiencia de haber depurado la base de datos de accidentalidad, se hace visible la tarea que tienen las autoridades de tránsito competentes, de realizar una actualización de protocolos e implementación de tecnologías al proceso de toma de datos en los eventos de accidentes viales. El uso de dispositivos móviles integrados con GPS y un almacenamiento centralizado eficientemente permitiría una rápida evaluación del comportamiento de siniestros en la ciudad por parte de los responsables de la movilidad.

La implementación de nuevos procedimientos no debe ser necesariamente costosa y ayudaría en gran medida al procesamiento rápido y análisis oportuno de toda la información recolectada. Lo ideal para la estandarización de estos procedimientos es que se hicieran y reglamentaran las políticas públicas desde el ámbito nacional, esto con el fin de lograr una homogenización en la calidad de

datos públicos del país, mejor aún, aplicando y discutiendo experiencias de nivel internacional.

En cuanto a los cuestionamientos que se han hecho sobre la geoestadística aplicada a información originaria de redes espaciales, se debe prestar atención a la validez de realizar estos análisis como se ha hecho, y posteriormente buscar métodos que se ajusten a la realidad de las redes y permitan la generación de resultados soportados en principios teóricos robustos.

Bibliografía

- Abdi, H. (2007). The Kendall Rank Correlation Coefficient. En *Encyclopedia of Measurement and Statistics*.
- Anderson, T. K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis and Prevention*, 41(3), 359-364. <https://doi.org/10.1016/j.aap.2008.12.014>
- Añez, J., De La Barra, T., & Pérez, B. (1996). Dual graph representation of transport networks. *Transportation Research Part B: Methodological*, 30(3), 209-216. [https://doi.org/10.1016/0191-2615\(95\)00024-0](https://doi.org/10.1016/0191-2615(95)00024-0)
- Batty, M. (2009). Accessibility: In Search of a Unified Theory. *Environment and Planning B: Planning and Design*, 36(2), 191-194. <https://doi.org/10.1068/b3602ed>
- Ben-Akiva, M., & Lerman, S. (1979). Disaggregate travel and mobility choice models and measures of accessibility. *Behavioural Travel Modelling*.
- Bulc, V. (2017). Speech by Commissioner Bulc at the ETSC Road Safety PIN Awards 2017. Recuperado 7 de julio de 2018, de https://ec.europa.eu/commission/commissioners/2014-2019/bulc/announcements/speech-commissioner-bulc-etsc-road-safety-pin-awards-2017_en
- Bunge, W. (1962). *Theoretical geography. Lund Studies in Geography Series C: General and Mathematical Geography*. Lund, Sweden.
- Burns, L. D. (1980). *Transportation, temporal and spatial components of accesibility*.
- Dalvi, M. Q., & Martin, K. M. (1976). The measurement of accessibility: Some preliminary results. *Transportation*. <https://doi.org/10.1007/BF00165245>
- DANE. (2016). ProyeccionMunicipios2005_2020. *Series de Población 1985 - 2020*. Recuperado de https://www.dane.gov.co/files/investigaciones/poblacion/proyepobla06_20/ProyeccionMunicipios2005_2020.xls
- De Fabritiis, C., Ragona, R., & Valenti, G. (2008). Traffic estimation and prediction based on real time floating car data. En *IEEE Conference on Intelligent Transportation*

- Systems, Proceedings, ITSC*. <https://doi.org/10.1109/ITSC.2008.4732534>
- Elvik, R. (2000). How much do road accidents cost the national economy? *Accident Analysis and Prevention*, 32, 849-851. Recuperado de http://ac.els-cdn.com/S0001457500000154/1-s2.0-S0001457500000154-main.pdf?_tid=53b5d16e-3caa-11e7-a0cd-00000aacb35e&acdnat=1495208970_82f2599b5d35ce12ff1dbb907ff7a050
- Escobar, D., & García, F. (2012). Territorial Accessibility Analysis as a Key Variable for Diagnosis of Urban Mobility: A Case Study Manizales (Colombia). *Procedia - Social and Behavioral Sciences*, 48, 1385-1394. <https://doi.org/10.1016/J.SBSPRO.2012.06.1114>
- Escobar, D., Martínez, S., & Moncada, C. (2016). Relación entre PM10 y condiciones de accesibilidad territorial urbana en manizales (Colombia). *Informacion Tecnologica*. <https://doi.org/10.4067/S0718-07642016000600027>
- Geurs, K. T., & van Eck, J. R. (2001). Accessibility measures: review and applications. Evaluation of accessibility impacts of land-use transportation scenarios, and related social and economic impact. *RIVM Report*.
- 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>
- Haig, B. D. (2004). What Is a Spurious Correlation? *Understanding Statistics*. https://doi.org/10.1207/s15328031us0202_03
- Haining, R. (2015). Spatial Autocorrelation. En *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*. <https://doi.org/10.1016/B978-0-08-097086-8.72056-3>
- Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, 25(2), 73-76. <https://doi.org/10.1080/01944365908978307>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). Kernel Smoothing Methods (pp. 191-218). Springer, New York, NY. https://doi.org/10.1007/978-0-387-84858-7_6
- Ingram, D. R. (1971). The concept of accessibility: A search for an operational form. *Regional Studies*, 5(2), 101-107. <https://doi.org/10.1080/09595237100185131>
- Instituto Nacional de Medicina Legal y Ciencias Forenses. (2018). *Forensis 2017: Datos para la Vida* (Vol. 19).

- Jones, A. (2016). Andrew Jones – 2016 Speech on Road Safety. Recuperado 27 de junio de 2018, de <http://www.ukpol.co.uk/andrew-jones-2016-speech-on-road-safety/>
- Jones, M. C. (1990). The performance of kernel density functions in kernel distribution function estimation. *Statistics & Probability Letters*, 9(2), 129-132. [https://doi.org/10.1016/0167-7152\(92\)90006-Q](https://doi.org/10.1016/0167-7152(92)90006-Q)
- Kenny, D. A. (1979). *Correlation and Causality*.
- Leduc, G. (2008). *Road Traffic Data: Collection Methods and Applications. Europe Communities*. <https://doi.org/JRC 47967 - 2008>
- Louf, R., & Barthelemy, M. (2013). Modeling the polycentric transition of cities. *Physical Review Letters*. <https://doi.org/10.1103/PhysRevLett.111.198702>
- Morales, P., & Rodríguez, L. (2016). *Aplicación de los coeficientes correlación de Kendall y Spearman*. Recuperado de <http://www.postgradovipi.50webs.com/archivos/agrollania/2016/agro8.pdf>
- Niroomand, N., & Jenkins, G. P. (2016). Estimating the Value of Life, Injury, and Travel Time Saved Using a Stated Preference Framework. *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2016.03.004>
- Organización Panamericana de la Salud (OPS). (2016). *La seguridad vial en la región de las américas. Organización Panamericana de la Salud*. <https://doi.org/10.2172/953225>
- Pardo Merino, A., & Ruiz Díaz, M. A. (2002). *SPSS 11: Guía para el análisis de datos*. McGraw-Hill. Recuperado de <https://www.casadellibro.com/libro-spss-11-guia-para-el-analisis-de-datos/9788448137502/861405>
- Rascón Chávez, O. A., Mendoza Díaz, A., & Mayoral Grajeda, E. F. (s. f.). *Tratamiento de sitios de alta incidencia de accidentes en carreteras mexicanas*. Recuperado de http://www.institutoivia.com/cisev-ponencias/analisis_accidentes_aa/Octavio_Agustin_Rascon.pdf
- Real Academia Española. (2001). Accidente. En *Diccionario de la lengua española (22.a ed.)*. Recuperado de <http://lema.rae.es/drae2001/srv/search?id=3EETCGquGDXX2nI8gwdT>
- Sheskin, D. (2000). *Handbook of Parametric and Nonparametric Statistical Procedures. Boca Raton: CRC*.
- Siegel, S., & Castellan, N. J. (1988). *Non-Parametric Statistics for the behavioural*

Sciences. *MacGraw Hill Int.*

- Sugiyanto, G. (2017). The cost of traffic accident and equivalent accident number in developing countries (Case study in Indonesia). *ARPJN Journal of Engineering and Applied Sciences*, 12(2), 389-397. Recuperado de http://www.arpnjournals.org/jeas/research_papers/rp_2017/jeas_0117_5631.pdf
- Sugiyanto, G., Fadli, A., & Santi, M. Y. (2017). Identification of black spot and equivalent accident number using upper control limit method, 12(2). Recuperado de www.arpnjournals.com
- Talen, E., & Anselin, L. (1998). Assessing spatial equity: An evaluation of measures of accessibility to public playgrounds. *Environment and Planning A*. <https://doi.org/10.1068/a300595>
- Tingvall, C., & Haworth, N. (1999). Vision Zero - An ethical approach to safety and mobility. En *Road Safety & Traffic Enforcement*.
- Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46. <https://doi.org/10.2307/143141>
- Wang, P., Hunter, T., Bayen, A. M., Schechtner, K., & González, M. C. (2012). Understanding road usage patterns in urban areas. *Scientific Reports*. <https://doi.org/10.1038/srep01001>
- Waters, N. (2018). Tobler's First Law of Geography. En *International Encyclopedia of Geography* (pp. 1-15). Oxford, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118786352.wbieg1011.pub2>
- World Health Organization. (2018a). *Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016*. Recuperado de http://www.who.int/gho/mortality_burden_disease/en/index.html
- World Health Organization. (2018b). *Global status report on road safety 2018*. Geneva. Recuperado de <http://e-journal.uajy.ac.id/14649/1/JURNAL.pdf>
- Xie, Z., & Yan, J. (2008). Kernel Density Estimation of traffic accidents in a network space. *Computers, Environment and Urban Systems*, 32(5), 396-406. <https://doi.org/10.1016/J.COMPENVURBSYS.2008.05.001>
- Yao, X. (2009). Georeferencing, Geocoding. *International Encyclopedia of Human Geography*, 458-465. <https://doi.org/10.1016/B978-008044910-4.00448-X>