



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA CIVIL

MODELAMIENTO ZONAL DE LA TASA DE ACCIDENTES DE TRÁNSITO EN SANTIAGO

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL

SERGIO ENRIQUE DROGUETT OVANDO

PROFESOR GUÍA:
ALEJANDRO TIRACHINI HERNÁNDEZ

MIEMBROS DE LA COMISIÓN:
MILTON BERTIN – JONES
LEONARDO BASSO SOTZ

SANTIAGO DE CHILE
JULIO 2018

RESUMEN DE LA MEMORIA PARA OPTAR AL

TÍTULO DE: Ingeniero Civil.

POR: Sergio Enrique Droguett Ovando

FECHA: 08/05/2018

PROFESOR GUÍA: Alejandro Tirachini Hernández

MODELAMIENTO ZONAL DE LA TASA DE ACCIDENTES DE TRÁNSITO EN SANTIAGO

El presente trabajo muestra una modelación de tasas de siniestralidad vial para la ciudad de Santiago mediante modelos lineales generalizados integrando tres formulaciones matemáticas distintas, buscando comparar los resultados y estimar un modelo que se ajuste lo mejor posible a la realidad y que tenga capacidad predictiva.

Dichas formulaciones corresponden a la regresión lineal Normal, regresión Poisson y regresión Binomial Negativa. Además, se utiliza el Método Empírico Bayesiano que ajusta de manera ponderada la modelación de tasas con el registro histórico de siniestros para considerar el fenómeno de regresión a la media y la sobredispersión. Además, se crean dos tipos de modelos: uno en donde las variables explicativas incluidas son significativas en todas las formulaciones, y otro donde estos regresores son estadísticamente significativos para al menos un tipo de regresión. Se utilizan variables que explican el fenómeno según lo observado en la literatura y según la disponibilidad de los datos existentes. Las fuentes de información son, principalmente, la base de datos de accidentes de tránsito de Carabineros de Chile, programas de operación del sistema de transporte público, información de bienes raíces del Servicio de Impuestos Internos y la encuesta de movilidad Origen-Destino del año 2012 del Ministerio de Transporte y Telecomunicaciones.

Se consideran cinco variables predichas, correspondientes a la tasa de accidentes totales, accidentes tipo no atropello, accidentes tipo atropello, cantidad de lesionados y cantidad de fallecidos. Lo anterior con una segregación espacial a nivel de zona obtenida de la Encuesta Origen-Destino del año 2012.

Se reportan estadísticos de bondad de ajuste que permiten comparar las estimaciones, junto con la predicción de accidentabilidad para el año 2015, tales como AIC y el error medio absoluto.

Los principales resultados de este trabajo indican que el mejor ajuste estadístico se realiza mediante la formulación Binomial Negativa utilizando el Método Empírico Bayesiano. Con respecto a los efectos marginales en el aumento de las tasas de siniestralidad vial, un peatón más será menos riesgoso que una persona más viajando en vehículo o en buses del transporte público.

DEDICATORIA

Especialmente dedicada a mis padres.

A mentores y mentoras de mi proceso formativo de ingeniero.

AGRADECIMIENTOS

Primero destaco el profundo agradecimiento a mi brillante profesor guía Alejandro Tirachini, quien con paciencia y sabiduría me acompañó en este trabajo que aborda una temática poco común en nuestro departamento. Sin duda él es el futuro de la seguridad vial y su relación con la academia en Chile.

Sin la motivación y el esfuerzo de mi familia, jamás esto podría haber sido alcanzado. Para Liliana Ovando Montoya y Sergio Droguett Morales mis gratitudes eternas.

Agradezco la disposición de los profesores Gabriel Montero, Luis Rizzi, Leonardo Basso y Milton Bertin-Jones por participar de la comisión revisora de este proyecto. Es para mí un honor tremendo contar con vuestro apoyo. Sinceros agradecimientos a los profesores Tomás Cox, Ricardo Hurtubia, Francisco Fresard, Mauro Huenupi, Marcela Munizaga, Francisco Martínez, Juan Enrique Cannobbio y muchos otros que son la base fundamental de la ingeniería en Chile y de quienes he tenido el orgullo de aprender.

Doy las gracias a mis compañeros entrañables de universidad y de la vida. Valoro en particular a Nicolás Maluenda y Leonardo Camus pues sus contribuciones fueron esenciales en este trabajo. Mención especial a Álvaro Fuentes, César Núñez, Elsa Durán, Alejandro Cárdenas, Juan Fernando Albornoz, Alejandro Gaitán, Daniela Arancibia, Elizabeth Contreras, Paz Bahamonde, Diego Saavedra, Gonzalo Ubilla, Patricio Wolff, Franco Basso y muchos otros que atesoro al finalizar este ciclo.

A mi pareja, Denisse Vilches, por su soporte incondicional y amor perpetuo.

A Raúl Espinoza y Mónica Chávez por su voluntad inquebrantable.

Estimo y aplaudo a todo aquel que dedica su trabajo en beneficio de la seguridad vial. Espero con ansias el día en que no se lamenten fatalidades en este contexto.

Finalmente, extendiendo un cariñoso saludo y agradecimiento a quienes olvido involuntariamente. Espero me perdonen.

TABLA DE CONTENIDO

RESUMEN	i
DEDICATORIA.....	ii
AGRADECIMIENTOS	iii
TABLA DE CONTENIDO.....	iv
ÍNDICE DE FIGURAS	vii
ÍNDICE DE GRÁFICOS	viii
ÍNDICE DE ECUACIONES.....	ix
ÍNDICE DE TABLAS	x
1 Introducción.....	1
2 Objetivos	3
3 Revisión Bibliográfica	4
3.1 Generalidades en el modelamiento del número de accidentes viales.....	4
3.2 Artículos científicos de modelamiento de accidentes y variables utilizadas	5
3.2.1 Variables Explicadas	6
3.2.2 Variables de Operación	6
3.2.3 Variables de Infraestructura.....	7
3.2.4 Variables de Regulación.....	7
3.2.5 Variables de Recopilación Post-Accidente	7
3.3 Modelos estadísticos utilizados.....	10
3.3.1 Modelos lineales generalizados.....	10
3.3.2 Familia Exponencial.....	10
3.3.3 Las componentes de un GLM.....	12
3.3.4 Método Empírico Bayesiano	13
3.4 Especificación de la formulación matemática.....	14
3.5 Ajuste de los modelos	16
3.5.1 Evaluación del ajuste.....	17
3.5.1.1 Prueba de Wald	17
3.5.1.2 Devianza	17
3.5.1.3 Criterio de información Akaike (AIC).....	17
3.5.1.4 Logaritmo de verosimilitud	18
3.5.1.5 Prueba ómnibus	18
3.5.1.6 Mean Absolute Error	18
4 Metodología	19

5	Recopilación de datos	20
5.1	Base de Datos SIEC2	20
5.2	Encuesta Origen-Destino del año 2012 para Santiago	26
5.2.1	Procesamiento de los datos	32
5.3	Base de Datos de Bienes Raíces del SII.....	35
5.4	Programas de Operación del Transantiago.....	39
5.5	Modelo de simulación basada en agentes a gran escala.....	43
5.6	Aforos DICTUC	45
6	Modelación.....	47
6.1	Variables explicadas	47
6.2	Variables explicativas.....	47
6.3	Modelos lineales generalizados Normal con función de enlace logarítmica.....	53
6.4	Modelos lineales generalizados Poisson con función de enlace logarítmica	58
6.5	Modelos lineales generalizados Binomial Negativa con función de enlace logarítmica	64
6.6	Análisis de las variables en los modelos	70
6.7	Análisis de efecto marginal sobre las tasas de accidentes de una persona más en modo auto, bus o caminata.....	76
6.8	Estimación aplicando el método empírico bayesiano.....	81
6.9	Comparación entre formulaciones	82
7	Estimación de tasas de accidentabilidad para el año 2015.....	95
7.1	Variaciones en las variables independientes	95
7.2	Comparación de lo predicho con lo observado	99
8	Conclusiones.....	110
8.1	Descripción del trabajo realizado	110
8.2	Resultados	111
8.3	Desafíos Futuros.....	114
9	Bibliografía	116
10	Anexos.....	118
10.1	Tablas Base SIEC2 y Descripción de los Campos	118
10.2	Distribuciones de accidentes geolocalizados por año.....	120
10.3	Descripción de la Base de Datos de Bienes raíces del SII	123
10.4	Algoritmo de rutas mínimas para datos OD de la EOD 2012.....	124
10.4.1	Generación de la Red.....	124

10.4.2	Generación de Tablas con Nodos Origen y Destino	124
10.5	Artículos Científicos con resumen de variables explicadas y explicativas ...	125
10.6	Estadísticos descriptivos de las variables	128
10.7	Pestañas de la modelación lineal generalizada en SPSS.....	130

ÍNDICE DE FIGURAS

Figura 1: Mapa de calor de los accidentes de tránsito de Santiago. Fuente: Elaboración propia en base a SIEC-2.	23
Figura 2: Cantidad de accidentes por zona EOD 2012. Fuente: Elaboración propia.	25
Figura 3: Zonificación considerada en la modelación y zonificación de la Encuesta Origen Destino 2012. Fuente: EOD 2012, Biblioteca del Congreso Nacional y elaboración propia.	27
Figura 4: Mapa de calor de los hogares considerados en la EOD 2012. Fuente: Elaboración propia en base a EOD 2012.	30
Figura 5: Mapas de calor de los orígenes (izquierda) y destino (derecha) de los viajes. Fuente: Elaboración propia en base a EOD 2012.	30
Figura 6: Mapa de calor de los viajes generados en caminata. Fuente: Elaboración propia en base a EOD 2012.	31
Figura 7: Mapa de calor de los viajes en bicicleta. Fuente: Elaboración propia en base a EOD 2012.	31
Figura 8: Mapa de la información contenida en el archivo Shapefile de bienes raíces. Fuente: Elaboración propia en base a datos facilitados por Tomas Cox y Ricardo Hurtubia.	38
Figura 9: Cobertura de las líneas de Transporte público, programa de operación primer semestre de 2013. Fuente: DTPM y elaboración propia.	39
Figura 10: Cantidad de buses promedio por día para cada zona. Fuente: Elaboración propia en base a datos DTPM.	41
Figura 11: Arcos de flujo vehicular. Fuente: Elaboración propia en base a tesis de Leonardo Camus.	43
Figura 12: Carga de flujo vehicular. Fuente: Elaboración propia en base a tesis de Leonardo Camus.	44
Figura 13: Flujo vehicular con densidad por arco. Fuente: Elaboración propia en base a tesis de Leonardo Camus.	44
Figura 14: Flujo de buses por zona para el año 2013 (izquierda) y para el año 2015 (derecha). Fuente: Elaboración propia.	96
Figura 15: Cantidad de accidentes totales en el año 2013 (izquierda) y 2015 (derecha). Fuente: Elaboración propia.	97
Figura 16: Distribución de accidentes en la región metropolitana para el año 2011. Fuente: Elaboración propia en base a datos SIEC-2 y BCN.	120
Figura 17: Distribución de accidentes en la región metropolitana para el año 2012. Fuente: Elaboración propia en base a datos SIEC-2 y BCN.	120
Figura 18: Distribución de accidentes en la región metropolitana para el año 2013. Fuente: Elaboración propia en base a datos SIEC-2 y BCN.	121
Figura 19: Distribución de accidentes en la región metropolitana para el año 2014. Fuente: Elaboración propia en base a datos SIEC-2 y BCN.	121
Figura 20: Distribución de accidentes en la región metropolitana para el año 2015. Fuente: Elaboración propia en base a datos SIEC-2 y BCN.	122

Figura 21: Pestaña Tipo de modelo en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	130
Figura 22: Pestaña de Respuesta en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	130
Figura 23: Pestaña Predictores en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	131
Figura 24: Pestaña de Modelo en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	131
Figura 25: Pestaña de Estimación en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	132
Figura 26: Pestaña Estadísticos en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	132
Figura 27: Pestaña Medias marginales estimadas en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	133
Figura 28: Pestaña Guardar en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	133
Figura 29: Pestaña Exportar en SPSS para la modelación lineal generalizada. Fuente: Elaboración propia en base a SPSS.	134

ÍNDICE DE GRÁFICOS

Gráfico 1: Gráfico de cantidad de viajes generados por comuna, incluyendo la cantidad de habitantes. Fuente: Elaboración propia en base a EOD 2012 y datos demográficos del INE.	29
Gráfico 2: Histograma de la cantidad de accidentes totales. Fuente: Elaboración propia.	84
Gráfico 3: Histograma de la cantidad de accidentes tipo no atropello. Fuente: Elaboración propia.	84
Gráfico 4: Histograma de la cantidad de accidentes tipo atropello. Fuente: Elaboración propia.	84
Gráfico 5: Histograma de la cantidad de lesionados. Fuente: Elaboración propia.	84
Gráfico 6: Histograma de la cantidad de fallecidos. Fuente: Elaboración propia.	85
Gráfico 7: Estadístico MAE para la modelación 2013, tipo 1. Fuente: Elaboración propia.	93
Gráfico 8: Estadístico MAE para la modelación 2013, tipo 2. Fuente: Elaboración propia.	94
Gráfico 9: Histogramas de los residuos para la modelación de tasas de accidentes totales del año 2015. Fuente: Elaboración propia.	100
Gráfico 10: Histogramas de los residuos para la modelación de tasas de accidentes tipo no atropello del año 2015. Fuente: Elaboración propia.	101
Gráfico 11: Histogramas de los residuos para la modelación de tasas de accidentes tipo atropello del año 2015. Fuente: Elaboración propia.	102

Gráfico 12: Histogramas de los residuos para la modelación de lesionados del año 2015. Fuente: Elaboración propia.	103
Gráfico 13: Histogramas de los residuos para la modelación de fallecidos del año 2015. Fuente: Elaboración propia.	104
Gráfico 14: Histogramas de los residuos del método empírico bayesiano de los accidentes totales. Fuente: Elaboración propia.	105
Gráfico 15: Histogramas de los residuos del método empírico bayesiano de los accidentes no atropello. Fuente: Elaboración propia.	105
Gráfico 16: Histogramas de los residuos del método empírico bayesiano de los accidentes atropello. Fuente: Elaboración propia.	106
Gráfico 17: Histogramas de los residuos del método empírico bayesiano de los lesionados. Fuente: Elaboración propia.	106
Gráfico 18: Histogramas de los residuos del método empírico bayesiano de los fallecidos. Fuente: Elaboración propia.	106
Gráfico 19: Estadístico MAE para la predicción 2015 de los modelos tipo 1. Fuente: Elaboración propia.	108
Gráfico 20: Estadístico MAE para la predicción 2015 de los modelos tipo 2. Fuente: Elaboración propia.	109

ÍNDICE DE ECUACIONES

Ecuación 1: Función generalizada de densidad de una variable aleatoria. Fuente: Elaboración propia.	10
Ecuación 2: Forma canónica de la función densidad de la familia exponencial. Fuente: Elaboración propia.	10
Ecuación 3: Forma generalizada de la función de densidad con parámetro de escala constante. Fuente: Elaboración propia.	11
Ecuación 4: Forma generalizada de la función densidad de la distribución normal. Fuente: Elaboración propia.	11
Ecuación 5: Predictor lineal. Fuente: Elaboración propia.	12
Ecuación 6: Relación entre la función de enlace y el predictor lineal. Fuente: Elaboración propia.	12
Ecuación 7: Ponderador método empírico bayesiano. Fuente: Elaboración propia.	13
Ecuación 8: Estimación por el método empírico Bayesiano. Fuente: Elaboración propia.	13
Ecuación 9: Función de probabilidad Poisson. Fuente: Elaboración propia.	14
Ecuación 10: Forma funcional de la estimación del parámetro de Poisson (frecuencia de accidentes). Fuente: Elaboración propia.	14
Ecuación 11: Forma funcional extendida del parámetro de Poisson. Fuente: Elaboración propia.	15
Ecuación 12: Varianza de la frecuencia de accidentes asumiendo distribución de error gamma. Fuente: Elaboración propia.	15

Ecuación 13: Especificación matemática utilizada en los modelos predictivos por estimar. Fuente: Elaboración propia.	15
Ecuación 14: Transformación de la formulación matemática del modelo mediante función de enlace logaritmo natural. Fuente: Elaboración propia.	16
Ecuación 15 Devianza. Fuente: Elaboración propia.....	17
Ecuación 16: Fórmula AIC. Fuente: Elaboración propia.....	18
Ecuación 17: Fórmula MAE Fuente: Elaboración propia.	18

ÍNDICE DE TABLAS

Tabla 3-1: Reporte de fatalidades según institución. Fuente: Elaboración propia en base al trabajo de Luis Rizzi, Ministerio de Salud (MINSAL, 2017) y Carabineros de Chile (CONASET, 2017).....	8
Tabla 3-2:Características de las distribuciones Normal, Poisson y Binomial Negativa. Fuente: Elaboración propia.	12
Tabla 5-1: Cantidad de Accidentes Localizados en la base de datos SIEC-2 procesada por CONASET. Fuente: Elaboración propia en base a SIEC-2.....	21
Tabla 5-2: Campos de la base de datos de accidentes utilizada, derivada de la base de datos SIEC-2. Fuente: Elaboración propia en base a SIEC-2.....	21
Tabla 5-3: Total de accidentes para cada comuna. Fuente: Elaboración propia en base a SIEC-2.....	22
Tabla 5-4: Cantidad de Accidentes por Grupo de causa para todos los años. Fuente: Elaboración propia en base a SIEC-2.	24
Tabla 5-5: Cantidad de víctimas según gravedad de las lesiones, por año. Fuente: Elaboración propia en base a SIEC-2.	24
Tabla 5-6: Cantidad de víctimas según gravedad de las lesiones por agrupación del tipo de accidente. Fuente: Elaboración propia en base a SIEC-2.....	24
Tabla 5-7: Riesgo de salud y de tráfico para la región metropolitana por cada año del estudio. Fuente: Elaboración propia en base a datos del INE y SIEC-2	25
Tabla 5-8: Cantidad de viajes generados en un día laboral normal según modo agregado. Fuente: EOD 2012 y elaboración propia.	28
Tabla 5-9: Detalle de problemas y resultados de las rutas mínimas para caminata, bicicleta y moto. Fuente: Elaboración propia.....	33
Tabla 5-10: Carga de caminata, bicicleta y peatones por comuna, promedio por día. Fuente: Elaboración propia.	34
Tabla 5-11: Cantidad de predios por año en cada semestre. Fuente: Elaboración propia en base a datos SII.	35
Tabla 5-12: Cuantía de roles por cada destino de uso principal, para cada año y semestre Fuente: Elaboración propia en base a datos SII.....	36
Tabla 5-13: Cantidad de metros cuadrados por destino de uso principal del predio para cada año y semestre. Fuente: Elaboración propia en base a datos SII.	37

Tabla 5-14: Indicadores de Frecuencia ICF para un día completo según operadores del Transantiago en distintos años y trimestres. Fuente: Elaboración propia en base a informes DTPM.	42
Tabla 5-15: ICF por semestre según unidad de negocio. Fuente: Elaboración propia en base a informes DTPM.	42
Tabla 5-16: Resumen de datos para calcular la tasa de ocupación. Fuente: Elaboración propia en base a aforos DICTUC (SECTRA, 2013).	45
Tabla 5-17 :Categorías asociadas a los diferentes niveles de ocupación de buses. Fuente: (SECTRA, 2013)	45
Tabla 5-18: Factores de ocupación por tipo de bus agregado. Fuente: (SECTRA, 2013)	46
Tabla 6-1: Variables consideradas en la modelación y su descripción. Fuente: Elaboración propia.	48
Tabla 6-2:Correlaciones entre las variables. Fuente: Elaboración propia.	49
Tabla 6-3:Factor de inflación de la varianza (VIF) por cada modelación. Fuente: Elaboración propia en base a modelación.	52
Tabla 6-4: Resultados de la modelación Normal para accidentes totales. Fuente: Elaboración propia.	54
Tabla 6-5: Resultados de la modelación Normal para accidentes tipo no atropello. Fuente: Elaboración propia.	55
Tabla 6-6: Resultados de la modelación Normal para accidentes tipo atropello. Fuente: Elaboración propia.	56
Tabla 6-7: Resultados de la modelación Normal para lesionados. Fuente: Elaboración propia.	57
Tabla 6-8: Resultados de la modelación Normal para fallecidos. Fuente: Elaboración propia.	57
Tabla 6-9: Resumen de Media y Varianza para distintos tipos de accidentes y frecuencias por zona. Fuente: Elaboración propia en base a SIEC-2	58
Tabla 6-10: Resultados de la modelación Poisson para accidentes totales. Fuente: Elaboración propia.	59
Tabla 6-11: Resultados de la modelación Poisson para accidentes tipo no atropello. Fuente: Elaboración propia.	60
Tabla 6-12: Resultados de la modelación Poisson para accidentes tipo atropello. Fuente: Elaboración propia.	61
Tabla 6-13: Resultados de la modelación Poisson para lesionados. Fuente: Elaboración propia.	62
Tabla 6-14: Resultados de la modelación Poisson para fallecidos. Fuente: Elaboración propia.	63
Tabla 6-15: Resultados de la modelación Binomial Negativa para accidentes totales. Fuente: Elaboración propia.	64
Tabla 6-16: Resultados de la modelación Binomial Negativa para accidentes tipo no atropello. Fuente: Elaboración propia.	66
Tabla 6-17: Resultados de la modelación Binomial Negativa para accidentes tipo atropello. Fuente: Elaboración propia.	67

Tabla 6-18: Resultados de la modelación Binomial Negativa para lesionados. Fuente: Elaboración propia.	68
Tabla 6-19: Resultados de la modelación Binomial Negativa para fallecidos. Fuente: Elaboración propia.	69
Tabla 6-20: Signos de los estimadores para las variables independientes utilizadas. Fuente: Elaboración propia.	71
Tabla 6-21: Resumen de las variables explicativas que se utilizan en cada modelo. Fuente: Elaboración propia.	73
Tabla 6-22: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de accidentes totales. Fuente: Elaboración propia.	77
Tabla 6-23: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de accidentes no atropello. Fuente: Elaboración propia.	78
Tabla 6-24: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de accidentes atropello. Fuente: Elaboración propia.	79
Tabla 6-25: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de lesionados. Fuente: Elaboración propia.	80
Tabla 6-26: Sobredispersión para los distintos modelos Binomial Negativa. Fuente: Elaboración propia.	82
Tabla 6-27: Estadísticos de bondad de ajuste de los modelos estimados. Fuente: Elaboración propia.	83
Tabla 6-28: Resumen de la modelación para la variable de accidentes totales, tipo 1. Fuente: Elaboración propia.	85
Tabla 6-29: Resumen de la modelación para la variable de accidentes totales, tipo 2. Fuente: Elaboración propia.	86
Tabla 6-30: Resumen de la modelación para la variable de accidentes no atropello, tipo 1. Fuente: Elaboración propia.	87
Tabla 6-31: Resumen de la modelación para la variable de accidentes no atropello, tipo 2. Fuente: Elaboración propia.	88
Tabla 6-32: Resumen de la modelación para la variable de accidentes atropello, tipo 1. Fuente: Elaboración propia.	89
Tabla 6-33: Resumen de la modelación para la variable de accidentes atropello, tipo 2. Fuente: Elaboración propia.	90
Tabla 6-34: Resumen de la modelación para la variable de lesionados, tipo 1. Fuente: Elaboración propia.	91
Tabla 6-35: Resumen de la modelación para la variable de lesionados, tipo 2. Fuente: Elaboración propia.	91
Tabla 6-36: Resumen de la modelación para la variable de fallecidos, tipo 1. Fuente: Elaboración propia.	92
Tabla 6-37: Resumen de la modelación para la variable de fallecidos, tipo 2. Fuente: Elaboración propia.	92
Tabla 7-1: Evolución de la partición modal 2001 – 2012. Fuente: (SECTRA, 2014).	95

Tabla 7-2: Factor de ampliación de los flujos al año 2015 por modo. Fuente: Elaboración propia.	96
Tabla 7-3: Factor de amplificación por uso de suelo para el año 2015. Fuente: Elaboración propia.	97
Tabla 10-1: Estructura Tabla Accidentes-Base Siec-2. Fuente: Análisis y Definición de una Metodología para la Evaluación Social de Impactos de Proyectos sobre la Seguridad Vial en Rutas Interurbanas (SECTRA, 2007).	118
Tabla 10-2: Estructura Tabla Personas-Base Siec-2. Fuente: Análisis y Definición de una Metodología para la Evaluación Social de Impactos de Proyectos sobre la Seguridad Vial en Rutas Interurbanas (SECTRA, 2007).	119
Tabla 10-3: Estructura Tabla Vehículos-Base Siec-2. Fuente: Análisis y Definición de una Metodología para la Evaluación Social de Impactos de Proyectos sobre la Seguridad Vial en Rutas Interurbanas (SECTRA, 2007).	119
Tabla 10-4: Campos de la Base de Datos de Bienes Raíces BBRR para la Información Básica de Roles No Agrícolas. Fuente: SII.	123
Tabla 10-5: Campos de la Base de Datos de Bienes Raíces BBRR para la Información de Terrenos y Construcciones No Agrícolas. Fuente: SII.	123
Tabla 10-6: Resumen de variables explicadas y explicativas en MPA: caso tramos de vía. Fuente: Metodología para la Evaluación Social de la Reducción de Accidentes Urbanos (SECTRA, 2014).	125
Tabla 10-7: Resumen de variables explicadas y explicativas en MPA: caso intersecciones. Fuente: Metodología para la Evaluación Social de la Reducción de Accidentes Urbanos (SECTRA, 2014).	126
Tabla 10-8: Modelos de Rotonda. Fuente: Elaboración propia.	127
Tabla 10-9: Modelo de Red. Fuente: Elaboración propia.	127
Tabla 10-10: Estadísticos descriptivos de las variables consideradas. Fuente: Elaboración propia.	129

1 Introducción

En Chile fallecen más de 1.500 personas al año debido a accidentes de tránsito. Más de 60.000 siniestros viales anuales producen un costo social de más de USD\$300.000.000 (Yañez, 2011). Según la OMS, la tasa de mortalidad debido a accidentes de tránsito por cada 100.000 habitantes (2013) ubican a África en la peor posición con una tasa de 26,6, el Mediterráneo Oriental con 19,9, luego la tasa mundial con 17,4, Pacífico Occidental con 17,3, Asia Sudoriental con 17,0, América con 15,9 y finalmente Europa con 9,3 (WHO, 2015).

Como se observa, los accidentes del tráfico son un problema en la sociedad. Representan externalidades negativas del transporte. Producen fallecidos, lesionados, congestión, costos administrativos para el estado, pérdidas de capital humano, entre otras. Es importante entonces que las entidades tomen un rol activo en este tema, y por lo mismo, recientemente ha sido publicada la nueva Política Nacional de Seguridad de Tránsito, la cual actualiza la versión del año 1993, en donde se definen 5 áreas de acción y 51 medidas para aumentar la seguridad vial. Estos ejes estratégicos contemplan la gestión de la seguridad vial, vías y movilidad más seguras, vehículos más seguros, usuarios de las vías más seguros y respuesta tras los siniestros de tránsito. Algunas de las medidas a implementar que destacan en lo relativo a este trabajo, son la incorporación de auditorías de seguridad vial en las etapas de un proyecto vial, gestión de la velocidad, promoción de buenas prácticas y recomendaciones de seguridad de tránsito para tomar decisiones más informadas, mejora en la captura de información de los lesionados en el sitio del siniestro para una oportuna y adecuada atención médica, mejorar tiempos de cruces para otorgar más seguridad a los peatones adultos mayores, fiscalización de velocidad, educación vial, fortalecer institucionalidad en la gestión, entre otras.

Por otro lado, por cada homicidio que ocurre en nuestro país hay casi 3 fallecidos en accidentes de tránsito (Yañez, 2011), lo que muestra que la importancia de la seguridad vial y su gravedad no se ven reflejadas en los medios de comunicación, en la agenda política ni en la opinión de la sociedad, debido a que no son temas prioritarios.

Cabe destacar que más del 83% de los accidentes entre el año 2000 y el año 2016 sucedieron en el ámbito urbano, además de corresponder a más del 71% de los lesionados a nivel país (CONASET, 2017). Además, la mayor parte de la población se ve afectada por estos tipos de accidentes pues son en las urbes donde habitan más personas: los habitantes de la Región Metropolitana representan poco más del 40% de la población nacional y, más generalmente, la población nacional urbana es más del 87% (Instituto Nacional de Estadística, 2017).

Sin perjuicio de lo anterior, la Secretaría de Planificación de Transporte (SECTRA) del Ministerio de Transporte y Telecomunicaciones (MTT) elaboró el año 2007 una metodología de evaluación social de reducción de accidentes para el caso interurbano que se formalizó en el Sistema Nacional de Inversiones. En la actualidad, el MTT amplió

el horizonte al caso urbano mediante un nuevo estudio que terminó en metodología de evaluación social (SECTRA, 2014).

En este contexto, una de las aristas de la metodología propuesta, son los modelos predictivos de accidentes de tránsito (MPA) pues pasan a ser una herramienta importante del proceso. Llegado el momento de calcular beneficios versus costos, es necesario establecer el efecto cuantitativo de una medida de seguridad sobre la frecuencia de accidentes. Con esto es posible establecer precios sociales y enmarcar una metodología de evaluación social de proyectos a la realidad nacional urbana. Entonces, los modelos predictivos son importantes no solo para estimar tasas de accidentes, sino que se encuadran en un proceso de utilidad mayor que, por ejemplo, es evaluar y responder a los lineamientos de inversión en el país ajustados a la necesidad de incluir la seguridad vial en los proyectos de infraestructura vial urbana.

Los modelos matemáticos ocupados para elaborar manuales de evaluación en otras partes del mundo, como en Estados Unidos (AASHTO, 2010) o Nueva Zelanda (New Zealand Transport Agency, 2013), utilizan regresiones lineales generalizadas bajo el supuesto de que los accidentes siguen unas distribuciones Poisson. Otros consideran que los accidentes deben ser tratados mediante una distribución binomial negativa, que corresponde a un caso particular de Poisson. Por otro lado, hay quienes proponen el modelo Zero-Inflated que considera que, por ejemplo, una esquina donde no haya accidentes debe ser incluida en la modelación y no descartada por su falta de siniestros.

En esta memoria, se presentan una serie de modelos que buscan explicar el fenómeno de la seguridad vial a escala zonal. Estos prototipos constan de variables explicadas como los accidentes peatonales, los no peatonales, fallecidos y lesionados. Las variables explicativas de los siniestros abarcan la pertenencia a una determinada comuna, variables de operación como el flujo en diversos modos y otras variables de uso de suelo. Finalmente se concluirá al respecto de qué variables observadas explican estadísticamente la ocurrencia de accidentes de tránsito y que se ajustan de mejor forma a la realidad observada.

2 Objetivos

El objetivo principal de esta memoria de título es comparar diferentes modelos de regresión y encontrar variables que sean estadísticamente significativas en explicar los accidentes de tránsito en ámbito urbano para la ciudad de Santiago, mediante modelos predictivos y descriptivos, a nivel de zona definida en la Encuesta Origen Destino del año 2012 (SECTRA, 2014) para las 34 comunas del Gran Santiago. Los modelos zonales permitirán la inclusión de nuevas variables explicativas, no incorporadas previamente en modelos de accidentes de tránsito en Chile.

Los objetivos secundarios son los siguientes:

- Revisar y sistematizar el estado del arte y las últimas investigaciones sobre la modelación de accidentes de tránsito.
- Caracterizar la ciudad y sus componentes que definen la situación del tráfico y los accidentes.
- Analizar y compactar la información nacional sobre flujos y accidentes de tránsito mediante procesos de geolocalización y extracción de datos de diversas fuentes como la EOD 2012, BBRR, SIEC-2, tesis, entre otros.
- Probar distintas formulaciones matemáticas de los modelos, incluyendo distintas formas funcionales y variables que definan el espacio vial (considerando la sobredispersión y el exceso de ceros).
- Comprobar la capacidad descriptiva de los modelos
- Comprobar la capacidad predictiva de los modelos
- Estudiar la sensibilidad y el ajuste de los modelos.

3 Revisión Bibliográfica

3.1 Generalidades en el modelamiento del número de accidentes viales

En la literatura hay numerosos estudios que plantean modelos econométricos de predicción de accidentes de tránsito. Estos pretenden entender de mejor forma los factores que afectan la ocurrencia de accidentes de tránsito, apuntando a entender el problema de las frecuencias de los siniestros, en vez de hallar la causa y el efecto de cada accidente de manera particular.

Este último punto plantea la necesidad de definir un solo foco de análisis pues, teniendo las dos aristas para abordar el problema, el enfoque sobre la frecuencia de accidentes es más recurrente en la literatura especializada, por sobre el análisis causa-efecto de accidentes individuales, que requiere de información muy detallada y puntual. Otra razón por la que se resuelve este camino es debido a que también presenta una posibilidad de extensión a otra área de análisis que es acorde con las necesidades de los proyectos de tránsito en el país: la evaluación económica por reducción o aumento de la seguridad vial.

Los datos para este tipo de modelos deben tener ciertas características que presentan desafíos metodológicos y problemas de estructura en la información. Algunos de estos inconvenientes han sido identificados y abordados (Lord & Mannering, 2010). A continuación, se presentan los habitualmente mencionados:

- a) Sobre y sub-dispersión: La varianza de los datos de tasa de accidentes es distinta de la media. Si es mayor, existe sobredispersión y viceversa. Un supuesto de la distribución de Poisson es que la media sea igual a la varianza.
- b) Variables explicativas que varían en el tiempo: Hay ciertas variables que cambian en el tiempo y deben ser correctamente abordadas para utilizarlos en modelos de corte temporal agregado.
- c) Correlación espacial y temporal: Al desagregar sobre el corte temporal de un mismo lugar para evitar el problema de la variación en el tiempo de las variables, se observará correlación entre las observaciones. Luego, esta correlación afecta la precisión del parámetro a estimar. Además, la correlación espacial se refiere al efecto de incluir 2 lugares o zonas adyacentes que pueden compartir efectos no observados.
- d) Bajo valor de la media de la muestra y pequeño tamaño de la muestra: Ya que los accidentes son eventos de baja frecuencia, su media tendrá un valor bajo y además la muestra de estos será también pequeña.
- e) Correlación de la severidad y tipo de accidentes: Un accidente del tipo colisión lateral, es más probable de tener personas involucradas con consecuencias fatales o graves. Otros tipos de accidentes, como choque o colisión frontal, son menos probables de tener consecuencias severas. En esto radica la correlación.

- f) Sub-reporte: No todos los accidentes son reportados en las bases de datos. Por ejemplo, hay algunos que por su naturaleza de baja severidad no alcanzan a ser denunciados o son omitidos en el ingreso a las bases de información.
- g) Variables omitidas o endógenas: Se intenta algunas veces desarrollar modelos simples con pocas variables explicativas. Los modelos entonces tendrán sesgos estadísticos importantes por omisión, al igual que toda estimación estadística tradicional.
- h) Forma funcional: lineal o no lineal: La mayoría de los modelos reportados en la literatura asumen cierta linealidad en el efecto de las variables explicativas de la frecuencia de accidentes de tránsito. Hay algunos que sugieren formas no-lineales que se ajustan mejor a la realidad.
- i) Parámetros fijos: Debido a cierta heterogeneidad no observada se puede asumir erróneamente que los parámetros no variarán entre ciertos lugares.

3.2 Artículos científicos de modelamiento de accidentes y variables utilizadas

Los métodos para modelar y estimar frecuencias de accidentes son variados y prometen validez estadística si son correctamente abordados y acotados. Esto tiene que ver con los supuestos y resultados obtenidos en el ajuste de análisis. Todos tienen ventajas y desventajas reportadas en la literatura, que terminan definiendo el alcance del estudio al estimar los modelos. El detalle de estos responde de manera particular a no solo la consideración metodológica y forma funcional matemática, sino que además a los requerimientos de información y sus limitantes.

Las variables explicadas y explicativas más utilizadas en la literatura revisada son las resumidas en la Tabla 10-6, Tabla 10-7, Tabla 10-8 y Tabla 10-9 del anexo 10.5. En ellas se puede observar que hay algunas que se repiten más que otras, como los flujos por acceso según vía prioritaria o secundaria, flujo peatonal, número de pistas y condición de luminosidad. Cabe destacar que estas tablas se realizaron en base a los estudios que se citan en los apartados siguientes, donde no solo se detectaron modelos para las intersecciones, sino que también para tramos de vía, rotondas y modelos de red. Este trabajo fue realizado para el estudio “Metodología para la evaluación social de la reducción de accidentes urbanos” (SECTRA, 2014).

Las variables fueron agrupadas según su origen en variables explicadas, de operación, infraestructura, regulación y de recopilación post-accidente. Esta agrupación puede tener variables que estén en más de una categoría a nivel conceptual, pero se simplifica su asignación pues se pretende solo observar la frecuencia de inclusión en modelos predictivos. Luego, esta información da origen al requerimiento y definición de datos a utilizar, y delimitan a grandes rasgos las variables más utilizadas que se observan en la literatura.

A continuación, se especifica cada nivel de variables y lo que se puede notar del análisis de la Tabla 10-6, Tabla 10-7, Tabla 10-8 y Tabla 10-9 en el anexo 10.5.

3.2.1 Variables Explicadas

Las variables explicadas que se observaron en la literatura tienen que ver con la cantidad de accidentes o frecuencia de éstos. Pueden o no incluir la severidad del siniestro, tipo de usuario, tipo de movimiento, por tipo de colisión o cantidad de vehículos involucrados.

Sobre esto es importante destacar que se consideran como variables explicadas en los modelos revisados las tasas de accidentes, por ejemplo, accidentes por cada habitante o por cada vehículo-kilómetro, y también la frecuencia o cantidad de accidentes por corte temporal. Pretenden ser útiles en el análisis por su aplicación directa sobre la probabilidad de accidentes en un espacio físico según variables observadas, en vez de considerar, por ejemplo, riesgos relativos.

Las variables más utilizadas en los modelos de las intersecciones fueron las de siniestros totales sin distinción (AASHTO, 2010; Gomes, 2013; Greibe, 2003; Layfield, Summersgill, Hall, & Chatterjee, 1996; M. J. Maher & Summersgill, 1996; Mountain, Fawaz, & Jarrett, 1996; New Zealand Transport Agency, 2013; Turner & Nicholson, 1998; Ye, Zhang, & Lord, 2013), seguida por la cantidad de accidentes por tipo de usuario y por tipo de movimiento (AASHTO, 2010; Layfield et al., 1996; M. J. Maher & Summersgill, 1996; Turner & Nicholson, 1998). El patrón de variables explicadas más utilizadas en los modelos por tramo de vía, rotonda y en red es el mismo: cantidad de accidentes totales sin distinción (AASHTO, 2010; Elvik, 2007; Gomes, 2013; Greibe, 2003; M. J. Maher & Summersgill, 1996; Mountain et al., 1996; New Zealand Transport Agency, 2013; Ye et al., 2013).

3.2.2 Variables de Operación

Las variables de operación tienen que ver con las características del tráfico en un espacio vial. Estas corresponden a los flujos tanto vehiculares como peatonales, sobre programación de semáforos, de circulación, de congestión, de diseño físico, entre otras.

Las variables de operación más utilizadas son las de flujo vehicular por acceso según vía prioritaria y secundaria (AASHTO, 2010; Elvik, 2007; Gomes, 2013; Greibe, 2003; Layfield et al., 1996; M. J. Maher & Summersgill, 1996; New Zealand Transport Agency, 2013; Turner & Nicholson, 1998; Ye et al., 2013), flujos vehiculares por tipo de movimiento o tipo de usuarios (Gomes, 2013; Layfield et al., 1996; M. J. Maher & Summersgill, 1996; New Zealand Transport Agency, 2013; Turner & Nicholson, 1998), flujo peatonal (Gomes, 2013; Layfield et al., 1996; M. J. Maher & Summersgill, 1996; New Zealand Transport Agency, 2013), velocidades de los vehículos (Layfield et al., 1996; M. J. Maher &

Summersgill, 1996; New Zealand Transport Agency, 2013) y presencia o número de pistas de viraje (Gomes, 2013; New Zealand Transport Agency, 2013).

3.2.3 Variables de Infraestructura

Las variables de infraestructura son aquellas que definen la vialidad a grandes rasgos en cuanto a las protecciones o facilidades de los usuarios en el espacio físico, y sobre el diseño vial que se utiliza.

Las variables más utilizadas son el número de pistas (AASHTO, 2010; Elvik, 2007; Gomes, 2013; Greibe, 2003; Layfield et al., 1996), largo de arco/segmento (AASHTO, 2010; Elvik, 2007; Gomes, 2013; M. J. Maher & Summersgill, 1996; Mountain et al., 1996; New Zealand Transport Agency, 2013; Ye et al., 2013), condición de luminosidad o visibilidad (AASHTO, 2010; Golob & Recker, 2004; Layfield et al., 1996; New Zealand Transport Agency, 2013) y el ancho de pista (Elvik, 2007; Gomes, 2013; Greibe, 2003; Layfield et al., 1996).

3.2.4 Variables de Regulación

Las variables de regulación tienen que ver con la legislación vigente en el lugar de estudio.

Las más frecuentes son las de velocidad límite (Elvik, 2007; Greibe, 2003; Layfield et al., 1996), regulación del cruce (AASHTO, 2010; Greibe, 2003; Layfield et al., 1996) y presencia de señalización (AASHTO, 2010; Layfield et al., 1996; M. J. Maher & Summersgill, 1996).

3.2.5 Variables de Recopilación Post-Accidente

Estas variables son aquellas que se recopilan luego de sucedido el siniestro vial. Son observadas y compiladas por entidades de seguridad pública y de interés privado, como por ejemplo Carabineros de Chile, aseguradoras o el Ministerio de Salud a través de su red asistencial en caso de lesionados o fallecidos.

Las variables más relevantes son el número y tipo de vehículos involucrados, edad de los ocupantes, condiciones climáticas cuando ocurre el accidente, tipo de accidente, cantidad de personas involucradas, gravedad y locación del siniestro (AASHTO, 2010; Carabineros de Chile, 2017; Elvik, 2007; Golob & Recker, 2004; Gomes, 2013; Greibe, 2003; Layfield et al., 1996; Lord, Washington, & Ivan, 2005; M. J. Maher & Summersgill, 1996; MINSAL, 2017; Mountain et al., 1996; New Zealand Transport Agency, 2013; Rizzi, 2001; Turner & Nicholson, 1998; Ye et al., 2013).

Uno de los enfoques más importantes en el análisis de los modelos predictivos de siniestralidad vial es el acceso a la información. Existen organizaciones internacionales dedicadas a recopilar información y a sugerir avances en esta materia. Uno de los ejemplos más importantes es International Traffic Safety Data and Analysis Group (IRTAD) del Foro Internacional de Transporte (ITF) perteneciente a la Organización para la Cooperación y el Desarrollo Económico (OECD).

El último informe de esta organización (IRTAD, 2017) describe y compara la gestión y el desempeño de la seguridad vial para diez países de América Latina, entre ellos, Chile. Entre las recomendaciones sobre la información y su manejo, se menciona adoptar una definición de 30 días posteriores a la ocurrencia de un siniestro vial en cuanto a la consideración de personas lesionadas o fallecidas. En Chile, la base de datos de accidentes viales reportada por Carabineros considera solo 24 horas. El sistema de salud reporta víctimas fatales que, independiente del lapsus de tiempo desde la ocurrencia del siniestro, mueren como consecuencia de este. En la base SIEC-2 no se realiza seguimiento particular de las víctimas para ajustar certeramente los datos. Esto se puede observar en la Tabla 3-1, donde se reportan entre un 23% y un 47% más fatalidades por parte del Departamento de Estadísticas e Información de Salud del Ministerio de Salud (MINSAL, 2017) versus lo reportado en la base de datos SIEC-2 de Carabineros de Chile (CONASET, 2017) entre los años 2000 y 2015 para todo el país.

Año	Carabineros de Chile	Ministerio de Salud	Diferencia Relativa
2000	1.698	2.197	1,29
2001	1.562	2.058	1,32
2002	1.549	2.018	1,30
2003	1.703	2.298	1,35
2004	1.757	2.311	1,32
2005	1.626	2.300	1,41
2006	1.652	2.309	1,40
2007	1.645	2.246	1,37
2008	1.782	2.476	1,39
2009	1.508	2.220	1,47
2010	1.595	2.170	1,36
2011	1.573	2.220	1,41
2012	1.523	2.129	1,40
2013	1.623	2.197	1,35
2014	1.630	2.065	1,27
2015	1.646	2.032	1,23
PROMEDIO TOTAL	1.630	2.203	1.35

Tabla 3-1: Reporte de fatalidades según institución.

Fuente: Elaboración propia en base al trabajo de Luis Rizzi, Ministerio de Salud (MINSAL, 2017) y Carabineros de Chile (CONASET, 2017).

Otro enfoque sobre el acceso a la información de la seguridad vial tiene que ver con las consideraciones de la experiencia internacional en cuanto al reporte de la causalidad. En un documento elaborado por académicos de la Pontificia Universidad Católica (Gazmuri et al, 2006) , se sugiere como medida complementaria para mejorar la seguridad vial del país, contar con mejoras constantes en la calidad de la información asociada a la

ocurrencia de cada accidente, en particular, sobre las causas, la ubicación, la hora, entre otras. Existen pautas internacionales al respecto que debiesen ser la base de dichas mejoras.

Todo lo anterior solo hace referencia a la información de los accidentes de tránsito, sin abordar la calidad de la información contenida en otras fuentes para el análisis de los factores de exposición y riesgo de la seguridad vial. En este sentido, es de vital importancia considerar que los datos sean lo más cercanos a la realidad, sin despreciar el hecho de que, al igual que toda base de datos, puede contener errores de todo tipo. Principalmente, estas faltas se pueden observar en el análisis posterior de los datos.

3.3 Modelos estadísticos utilizados

3.3.1 Modelos lineales generalizados

Los modelos lineales generalizados (GLM por sus siglas en inglés: *Generalized linear models*) proveen una unificación acerca de los procedimientos estadísticos más comunes, usados en estadística aplicada. Son una generalización flexible de la regresión lineal tradicional.

Relaciona la distribución aleatoria de la variable dependiente con la parte sistemática o no aleatorio a través de una función de enlace.

Los modelos lineales generalizados fueron formulados por John Nelder y Robert Wedderburn (Nelder & Wedderburn, 1972). Intentaban unificar varios modelos estadísticos, entre ellos la regresión lineal, regresión Poisson y regresión logística.

En estos modelos se asume que la variable dependiente está generada por una función de distribución de la familia exponencial.

El desarrollo presentado a continuación está fundado en lo descrito en el libro de James Lindsey (Lindsey, 2000).

3.3.2 Familia Exponencial

Muchas distribuciones estadísticas forman parte de un gran grupo llamado “familia exponencial”. Presentan una ventaja con respecto a otros modelos que se refiere al cálculo de los estadísticos.

Consideran una variable de respuesta aleatoria e independiente y_i ($i = 1, 2, 3, \dots, n$) y una función de probabilidad (densidad) de la forma mostrada en la Ecuación 1.

$$f(y_i, \theta_i) = r(y_i) * s(\theta_i) * \exp(t(y_i) * u(\theta_i)), \quad y_i \in \mathbb{R}$$

*Ecuación 1: Función generalizada de densidad de una variable aleatoria.
Fuente: Elaboración propia.*

Con θ_i un parámetro de localización indicando la posición donde la distribución está dentro del rango de posibles valores de respuesta. Cualquier distribución que pueda ser escrita de esta manera es parte de la familia exponencial con un parámetro.

Se puede considerar una forma canónica como la Ecuación 2.

$$f(z_i, \gamma_i) = \exp(z_i * \gamma_i + b(\gamma_i) + c(z_i))$$

*Ecuación 2: Forma canónica de la función densidad de la familia exponencial.
Fuente: Elaboración propia.*

Donde $z_i = f(y_i)$ y $\gamma_i = u(\theta_i)$ funciones de valor real y $b(\gamma_i)$ es la constante de normalidad de la distribución.

Las variables aleatorias y_i tendrían así una media μ_i , con lo que se tendría el modelo clásico estadístico $y_i = \mu_i + \varepsilon_i$.

Lo anterior puede ser aún más generalizado incluyendo un parámetro de escala (constante) φ_i como se muestra en la Ecuación 3.

$$f(z_i, \gamma_i, \varphi) = \exp\left(\frac{z_i * \gamma_i - b(\gamma_i)}{a_i(\varphi)} + c(z_i, \varphi)\right)$$

*Ecuación 3: Forma generalizada de la función de densidad con parámetro de escala constante.
Fuente: Elaboración propia.*

Donde γ_i es la forma canónica del parámetro de localización para alguna función de la media μ_i .

A modo de ejemplo, se puede notar la distribución normal como en la Ecuación 4.

$$\begin{aligned} f(z_i, \mu_i, \sigma^2) &= \frac{1}{\sqrt{2 * \pi * \sigma^2}} * \exp\left(-\frac{(z_i - \mu_i)^2}{2 * \sigma^2}\right) \\ &= \exp\left(\frac{\left(z_i * \mu_i - \frac{\mu_i^2}{2}\right)}{\sigma^2} - \frac{1}{2} * \left(\frac{z_i^2}{\sigma^2} + \log(2 * \pi * \sigma^2)\right)\right) \end{aligned}$$

*Ecuación 4: Forma generalizada de la función densidad de la distribución normal.
Fuente: Elaboración propia.*

Donde $\gamma_i = \mu_i$, $b(\gamma_i) = \frac{\mu_i^2}{2}$, $a_i(\varphi) = \sigma^2$ y $c(z_i, \varphi) = -\frac{z_i^2 + \log(2 * \pi * \sigma^2)}{2}$.

Cabe destacar que dentro de la familia exponencial se consideran las siguientes distribuciones, además de la normal antes mostrada:

- Exponencial
- Gamma
- Chi-cuadrado
- Bernoulli
- Poisson
- Binomial
- Binomial Negativa
- Geométrica.

No son parte de la familia las distribuciones de Cauchy y uniforme.

3.3.3 Las componentes de un GLM

Al considerar un modelo lineal generalizado (GLM) se deben tener en cuenta 3 componentes importantes de su definición.

En primer lugar, en la formulación de un GLM se considera la componente aleatoria independiente referida a $y_i (i = 1, 2, 3, \dots, n)$ con medias μ_i . En ellas se notará alguna distribución de la familia de distribución exponencial con un parámetro de escala constante.

Otra componente es la sistemática, que se refiere a un vector de variables explicativas. Bajo esta definición, dicho vector queda definido en la Ecuación 5.

$$\tau_i = \sum_j \beta_j * x_{ij} \quad i = 1, 2, 3, \dots, n \quad j = 1, 2, 3, \dots, p$$

*Ecuación 5: Predictor lineal.
Fuente: Elaboración propia.*

Con x_{ij} el predictor j para el sujeto i , β_j el coeficiente a estimar de las p variables explicativas. A esta combinación lineal de variables se le llama predictor lineal.

La última componente tiene que ver con la llamada función de enlace, en la que si se considera que $\gamma_i = \tau_i$, la relación entre la media y el predictor lineal será a través de una función de enlace, como se muestra en la Ecuación 6.

$$\tau_i = g_i(\mu_i) = \sum_j \beta_j * x_{ij} = \beta * X_i^T$$

*Ecuación 6: Relación entre la función de enlace y el predictor lineal.
Fuente: Elaboración propia.*

Esta función debe ser monótona y diferenciable. Además, esta función es aquella que transforma la media a el parámetro de localización canónico de la familia de distribución exponencial. Para este trabajo se utilizará en todas las formulaciones la función de enlace logarítmica debido a la especificación presentada en el apartado siguiente.

En la Tabla 3-2 se muestran las características más relevantes de las distribuciones Normal, Poisson y Binomial Negativa.

Distribución	Rango de y_i	Función de distribución $f(y)$	$\mu(\gamma)$	Varianza (μ)	$a(\varphi)$
Normal	$(-\infty, \infty)$	$\frac{\exp\left(-\frac{(y-\mu)^2}{2 * \sigma^2}\right)}{\sqrt{2 * \pi * \sigma^2}}$	γ	1	σ^2
Poisson	$[0, 1, 2, 3, \dots]$	$\frac{\mu^y}{y!} * \exp(-\mu)$	$\exp(\gamma)$	μ	1
Binomial Negativa	$[0, 1, 2, 3, \dots]$	$\binom{k+y-1}{y} * \left(\frac{\mu}{k+\mu}\right)^y * \left(\frac{k}{k+\mu}\right)$	$\frac{k * \exp(\gamma)}{1 - \exp(\gamma)}$	$\mu + \frac{\mu^2}{k}$	1

*Tabla 3-2: Características de las distribuciones Normal, Poisson y Binomial Negativa.
Fuente: Elaboración propia.*

3.3.4 Método Empírico Bayesiano

Buscando revisar la experiencia internacional en cuanto a la modelación de accidentes de tránsito, se tomó en cuenta el trabajo realizado en los países con alta experiencia en el ámbito de la seguridad vial. En este sentido es que se observó que se realizan métodos para estimar la accidentabilidad que consideran modelos predictivos y los datos históricos de los siniestros en los lugares de análisis. Este método se conoce como “empírico bayesiano” y representa el estado más avanzado en la práctica. Este tipo de procedimiento fue utilizado en los estudios realizados en Chile por la consultora Aristo por mandato de SECTRA (SECTRA, 2007, 2014).

El método empírico bayesiano dicta que la mejor estimación de los accidentes de tránsito es el promedio ponderado entre el valor obtenido por los modelos predictivos y los datos de accidentes registrados en el lugar de análisis. Dicho ponderador (α) es el presentado en la Ecuación 7.

$$\alpha = \frac{1}{1 + sd * acc_m}$$

Ecuación 7: Ponderador método empírico bayesiano.

Fuente: Elaboración propia.

Donde:

- sd : Es la sobredispersión.
- acc_m : Accidentes estimados por la modelación.

A sd se le conoce como el parámetro de la sobre dispersión. Este se refiere a cuando la media es menor que la varianza. Permite concluir que el modelo de Poisson es un caso particular del modelo Binomial Negativo, como se mostrará más adelante, pues corresponde al caso en que sd tiende a cero, es decir, no hay sobredispersión y por lo tanto la media (o esperanza) es igual a la varianza.

Luego, la estimación por el método empírico bayesiano es la presentada en la Ecuación 8.

$$acc = \alpha * acc_m + (1 - \alpha) * acc_r$$

Ecuación 8: Estimación por el método empírico Bayesiano.

Fuente: Elaboración propia.

Donde:

- acc_r : Accidentes registrados en el lugar de análisis según si son accidentes totales, accidentes peatonales, no peatonales, accidentes con lesionados o accidentes con fallecidos.
- acc : Accidentes estimados por el método empírico Bayesiano.

Lo anterior muestra que en una estimación donde no se toma en cuenta la sobredispersión, como por ejemplo un modelo lineal o una regresión Poisson, se estarán

sesgando los efectos de los accidentes registrados debido a que la modelación sesga los coeficientes estimados y considera que $\alpha = 1$.

En la literatura también se reporta el fenómeno de la regresión a la media. Este consiste en que, para un cierto fenómeno en observación, existen factores probabilísticos que pueden hacer que los valores observados se alejen de la media, ya sea inferior o superiormente, de manera transitoria. Es particularmente relevante en el análisis de la seguridad vial debido a que, en general, cuando ocurre un hecho significativo (por ejemplo, un accidente con varias personas fallecidas) suelen producirse cambios en la legislación o en la percepción y fiscalización por parte de las autoridades. Esto conllevaría a que la futura estimación de la mejora en seguridad vial tenga asociado un sesgo importante debido a la regresión a la media cuando no se constata que el efecto de un alza o una baja en los datos es debido a variaciones naturales de los datos (M. Maher & Mountain, 2009).

El método empírico bayesiano considerará siempre que los accidentes será un valor intermedio entre lo estimado y lo observado, y, por lo tanto, considerará el efecto de regresión a la media.

3.4 Especificación de la formulación matemática

La especificación matemática de los modelos de accidentes de tránsito más utilizada y que ha marcado los comienzos de los modelos predictivos es la regresión Poisson. Considera que la probabilidad de ocurrencia de un accidente distribuye de manera Poisson con tasa igual a la frecuencia de accidentes. Esta especificación establece que los accidentes son eventos de baja ocurrencia y su frecuencia es una variable entera no negativa.

Este modelo de regresión asume que una vía i (segmento, intersección, zona, comuna, etc.) tiene probabilidad de accidentes por cierta unidad de tiempo y_i , que viene dada por la fórmula en la Ecuación 9.

$$P_n(y_i|\mu_i) = \frac{(\mu_i)^{y_i} * e^{-\mu_i}}{y_i!}$$

*Ecuación 9: Función de probabilidad Poisson.
Fuente: Elaboración propia.*

Donde $P_n(y_i)$ es la probabilidad de que en la locación i ocurran y_i accidentes por unidad de tiempo, y μ_i es el parámetro de Poisson o, que es lo mismo, la frecuencia esperada de accidentes de algún tipo por unidad de tiempo. Luego, esta última se debe estimar en función de variables explicativas de la forma funcional en la Ecuación 10.

$$\mu_i = e^{X_i\beta}$$

*Ecuación 10: Forma funcional de la estimación del parámetro de Poisson (frecuencia de accidentes).
Fuente: Elaboración propia.*

Donde X_i es un vector de variables características del lugar i , y β el vector de parámetros a estimar.

A pesar de que este modelo sirvió de punto de partida para muchos investigadores en el análisis de la frecuencia de los accidentes, el supuesto que estos siguen una distribución Poisson asume que su media debe ser igual a su varianza. En muchos casos se ha visto que esto no es cierto para los siniestros viales, en donde se produce sobredispersión (Lord & Mannering, 2010). Además, la robustez del modelo se ve afectada cuando se tienen valores bajos de la media y cuando el tamaño muestral es pequeño.

Una sofisticación de éste es el modelo Binomial Negativo o Poisson-Gamma, el cual sigue asumiendo que los eventos distribuyen Poisson, pero el parámetro λ_i ahora se estima según la fórmula en la Ecuación 11.

$$\mu_i = e^{X_i\beta + \delta_i}$$

*Ecuación 11: Forma funcional extendida del parámetro de Poisson.
Fuente: Elaboración propia.*

Donde e^{δ_i} es un error que distribuye gamma de media 1 y varianza sd . Ahora la varianza difiere de la media según la Ecuación 12.

$$Var[y_i] = E[y_i] * (1 + sd * E[y_i])$$

*Ecuación 12: Varianza de la frecuencia de accidentes asumiendo distribución de error gamma.
Fuente: Elaboración propia.*

Una consideración importante que se destaca de la literatura es que si bien la especificación de la Ecuación 10 puede considerar las variables explicativas en el vector X_i , existe una mejora en cuanto a la consideración de los efectos de flujos en diversos modos, transformándolos a una forma potencial. Esto da cabida a la formulación utilizada en este trabajo, presentada en la Ecuación 13.

$$\mu_i = \left(\prod_m F_{im}^{\beta_m} \right) * e^{\sum_j \beta_j * x_{ij}}$$

*Ecuación 13: Especificación matemática utilizada en los modelos predictivos por estimar.
Fuente: Elaboración propia.*

Donde m representa los diversos modos considerados (caminata, vehículo particular, transporte público, bicicleta, entre otros), F_m son los flujos para dichos modos, y β_m y β_j son los coeficientes por estimar para los m modos y las j variables explicativas x_{ij} . Debido a esta especificación, se destaca que el estimador de los flujos en los diversos modos correspondería a la elasticidad del flujo versus la tasa de accidentalidad. Es decir, un aumento en un 1% del flujo, debiese producir un aumento en $\beta_m\%$ la tasa de accidentes.

Cabe destacar que, considerando una función de enlace logaritmo natural, la especificación se puede transformar en la Ecuación 14.

$$\ln(\mu_i) = \sum_m \beta_m * \ln(F_{im}) + \sum_j \beta_j * x_{ij}$$

*Ecuación 14: Transformación de la formulación matemática del modelo mediante función de enlace logaritmo natural.
Fuente: Elaboración propia.*

Existen más sofisticaciones del modelo de regresión de Poisson, como el Poisson-log normal, que asume que el término del error ahora sigue una distribución log-normal.

El caso más restrictivo para los modelos se da cuando existen muchos ceros en una locación i . Este suceso es abordado por el modelo “Zero-Inflated”, que opera bajo el principio que el exceso de ceros se puede acomodar al separar la modelación entre la propensión de un lugar de tener accidentes y su inclinación por no tenerlos. La probabilidad de una vía a estar en el estado cero o no-cero puede ser determinado por un modelo binario Logit o un Probit. El caso no-cero se puede estimar por el supuesto de que el proceso se rige por el modelo Poisson o por el modelo Binomial Negativo. De esto último es que se apellida al modelo Zero Inflated Poisson o Negative Binomial.

Este modelo ha sido uno de los más populares en el último tiempo, y a su vez ha recibido críticas que hablan sobre su incapacidad de reflejar el proceso de generación de datos cuando el estado “cero accidentes” tiene una media de largo plazo igual a cero (Lord & Mannering, 2010).

En la literatura se notan muchos otros modelos, donde cada uno pretende abordar de la mejor manera posible la adversidad en los datos mostrado anteriormente. Algunos de estos son: Conway-Maxwell-Poisson, Generalized estimating equation, Generalized additive, Random-effects, Negative multinomial, Random-parameters, Finite mixture/Markov switching, Hierarchical/multilevel, Neural network, Bayesian neural network and support vector machine, entre otros (Lord & Mannering, 2010).

La modelación presentada en este trabajo considera las formulaciones lineales, de Poisson y Binomial Negativa, de manera de abordar el problema comparando los resultados para diversos tipos de variable explicada.

En el desarrollo de esta memoria de título, se hará uso de la herramienta computacional IBM SPSS con paquetes estadísticos programados sobre estas formulaciones para estimar los modelos. Principalmente se hará uso de las herramientas de regresión lineal y de regresión lineal generalizada.

3.5 Ajuste de los modelos

En la regresión lineal, el método más utilizado para estimar los parámetros es el de mínimos cuadrados ordinarios. El método general de la estimación que conduce a la función de mínimos cuadrados se llama método de máxima verosimilitud.

Este último método produce valores de los parámetros que maximizan la probabilidad de obtener un conjunto de datos observados. Para la aplicación del método se debe construir

la función de verosimilitud que expresa la probabilidad de los datos observados como una función de los parámetros desconocidos y luego obtener su maximización en torno al parámetro.

3.5.1 Evaluación del ajuste

Para la evaluación del ajuste se consideran diversos estadígrafos que son utilizados en el análisis del ajuste de los GLM.

3.5.1.1 Prueba de Wald

La prueba de Wald o el test chi cuadrado de Wald es una manera de determinar si las variables explicativas en un modelo son significativas o no. La hipótesis nula con la que se contrasta es si el coeficiente de la variable podría ser cero, versus su hipótesis alternativa que es que sea significativamente distinta de cero. Se compara con la distribución χ^2 y de ahí su nombre. El p-valor o significancia del test indica si el coeficiente es significativamente distinto de cero para un nivel de confianza de 95% cuando dicho valor sea inferior a 0,05.

El test de Wald es una aproximación de la prueba de razón de verosimilitud. Para valores de la muestra grandes es equivalente al test t. Su uso no es recomendado para tamaños de muestras menores a 30 observaciones.

3.5.1.2 Devianza

Se entiende como devianza a la cantidad de varianza explicada por el modelo. Da una idea de la variabilidad explicada de los datos al compararla con el modelo nulo. Su fórmula se presenta en la Ecuación 15.

$$D^2 = \frac{\text{Devianza modelo nulo} - \text{Devianza residual}}{\text{Devianza modelo nulo}}$$

*Ecuación 15 Devianza.
Fuente: Elaboración propia*

3.5.1.3 Criterio de información Akaike (AIC)

Es criterio de información de Akaike (AIC) es una medida de la calidad relativa de un modelo estadístico entre la bondad de ajuste y la complejidad del modelo. Se basa en la entropía de la información y se debe notar que no proporciona una prueba de hipótesis, por lo que AIC no indica nada acerca de la calidad del modelo en un sentido absoluto,

sino que representa solo una medida relativa en la que se pueden preferir siempre valores menores de AIC ya que representa mejores ajustes.

Se calcula con la fórmula de la Ecuación 16.

$$AIC = -2 * \left(\frac{l}{N}\right) + 2 * \left(\frac{k}{N}\right)$$

*Ecuación 16: Fórmula AIC.
Fuente: Elaboración propia.*

Donde l es el logaritmo de la función verosimilitud, N es el número de observaciones y k es la cantidad de parámetros estimados.

3.5.1.4 Logaritmo de verosimilitud

Es la función de verosimilitud la que se busca maximizar. Por lo mismo, a mayor valor del logaritmo de verosimilitud o *log likelihood*, se obtendrá un mejor ajuste. Sirve también para calcular estadísticos que indiquen el ajuste del modelo.

3.5.1.5 Prueba ómnibus

La prueba de contraste ómnibus es un test de chi-cuadrado de la razón de verosimilitud del modelo actual frente al modelo de interceptación. Si el valor de significancia es inferior a 0,05, el modelo actual funciona mejor que el modelo nulo.

3.5.1.6 Mean Absolute Error

El error absoluto medio (o Mean Absolute Error MAE en inglés) es un promedio de diferencia entre las observaciones predichas y las observadas. Su fórmula es la mostrada en la Ecuación XXX.

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

*Ecuación 17: Fórmula MAE
Fuente: Elaboración propia.*

Donde y_i es la variable explicada observada en la realidad, \hat{y}_i es la variable explicada estimada por los modelos y n es la cantidad de observaciones, que en este trabajo corresponde a la cantidad de zonas.

4 Metodología

Para desarrollar los objetivos de este estudio, es necesario recopilar la información de diversas fuentes que puedan ser significativas en el análisis de los modelos.

Estas fuentes deben considerar información relevante a la generación y atracción de viajes, así como también del modo en que se viaja y los niveles de flujo existentes en la ciudad. Los datos consideran la experiencia internacional en el modelamiento de este fenómeno, además de tener presente que corresponden a un proxy a nivel macro de los factores que puedan influir en la ocurrencia de un siniestro de tránsito.

Por otro lado, es necesario que esta información esté georreferenciada de tal manera de poder ubicar en el espacio tanto los accidentes de tránsito a nivel de zona como la información recopilada. Esto considera un importante desafío al momento de conseguir la información y consta de un proceso costoso en el desarrollo de los modelos.

Un sistema de información geográfica (GIS por sus siglas en inglés) es un conjunto de herramientas que permiten la organización, almacenamiento, manipulación, análisis y modelización de grandes cantidades de datos procedentes del mundo real que están vinculados a una referencia espacial. Más genéricamente, los GIS son herramientas que permiten a los usuarios crear consultas interactivas, analizar información espacial geográficamente referenciada, editar datos, mapas y presentar los resultados de todas las operaciones en cierta representación de datos georreferenciados.

Tanto los accidentes como la mayoría de la información contenida en diversas fuentes están georreferenciados, por lo que es necesaria la utilización de un software capaz de ofrecer herramientas de análisis. En esta memoria de título se trabaja con el software ArcGis a través de su licencia de prueba para estudiantes y con el software QGIS de código libre, además de las funcionalidades del sistema de administración de base de datos Postgres y sus extensiones PostGIS y pgrouting.

Al comienzo de este trabajo se consideró trabajar con datos a nivel de lugar, como por ejemplo en una intersección o tramo de vía, tal cual lo hacen en los estudios de metodología de evaluación social de accidentes de tránsito realizados por la consultora Aristo (SECTRA, 2014). Sin embargo, el proceso fue modificado para ser elaborado con los datos disponibles de la EOD 2012 debido a que con ellos se cuenta con un nivel de detalle significativo y permite la incorporación de nuevas variables explicativas, como el flujo de buses y de peatones, además de que esos resultados fueron publicados el año 2015, posterior al desarrollo de los modelos predictivos de accidentes de tránsito realizados en dichos estudios de la consultora.

Finalmente, con toda la información compactada y corregida, se procede a realizar el modelamiento según lo especificado. Los modelos serán analizados en cuanto a su composición, robustez, significancia y nivel de precisión al describir la realidad de los siniestros viales.

5 Recopilación de datos

Con el objetivo de contar con datos que caractericen el lugar o zona de ocurrencia de un accidente, se intentaron varias formas de acopiar información que se observa relevante según los estudios internacionales. En los apartados siguientes se muestra el resultado de dicho proceso según las diferentes fuentes consultadas.

5.1 Base de Datos SIEC2

Carabineros de Chile registra los accidentes de tránsito en terreno, cuya información luego es recopilada para formar la base de datos del Sistema Integrado Estadístico de Accidentes de Tránsito y Ferroviarios de Carabineros de Chile, formulario 2 (SIEC-2). Además, la Comisión Nacional de Seguridad de Tránsito (CONASET) del Ministerio de Transporte y Telecomunicaciones (MTT) mantiene acuerdos de trabajo en los que ejecuta diversas tareas sobre la información. Estos datos son relevantes dado que con ellos se podrá modelar la variable explicada en los Modelos de Predicción de Accidentes (MPA) del estudio (SECTRA, 2014) y de este trabajo.

Consta de las tablas “Accidentes”, “Personas” y Vehículos”, pudiendo identificarse las variables incluidas en el anexo 10.1. Para la ciudad de Santiago, una de las tareas que realiza CONASET es procesar la información de geolocalización de casi todos los siniestros de 34 comunas del Gran Santiago.

En (SECTRA, 2014) se constatan varios problemas a tener en cuenta al trabajar con estos datos, las cuales han sido críticas recurrentes entre los expertos académicos durante ya un buen tiempo, como se menciona en el apartado de referencia bibliográfica. Uno de estos problemas es el subreporte de los accidentes sin lesionados, a pesar de que se observa que son la mayoría. Por otro lado, la localización de los accidentes no siempre está determinada o reportada, lo que hace subestimar el número de accidentes existentes. Además, la causa del accidente es categorizada de manera única, siendo que la buena práctica internacional dicta que, al ser un fenómeno multifactorial, deben ser varias las procedencias reportadas y ajustadas. Finalmente, los registros no siempre cuentan con información completa, los campos no siempre permiten realizar análisis detallados de cada siniestro o se repiten los registros.

En el trabajo de la presente memoria se cuenta con información facilitada por CONASET a través de una solicitud por ley de Transparencia (MINSEGPRES, 2008) con datos ya depurados. En particular, la información de localización ya fue georreferenciada y se cuenta con la ubicación de los accidentes. Sin embargo, se realizaron de igual forma análisis de consistencia, información incompleta, normalización de algunos valores, entre otros. Los años con los que se tiene información son el 2011, 2012, 2013 y 2014.

Cabe destacar que la mencionada base está en formato Shapefile, el cual es un formato de archivo informático con información geográfica vectorizada de ciertos elementos y

atributos asociados a ellos. Se puede manejar con softwares de Sistemas de Información Geográfica (SIG) como ArcGis, QGIS o NEXTGIS.

Los accidentes localizados en las 34 comunas de Santiago, que no fueron georreferenciados por CONASET se muestran en la Tabla 5-1. Equivalen a 8,28% del total.

Año	Accidentes Localizados	No ubicados
2011	17.572	3.112
2012	17.150	2.632
2013	18.379	1.604
2014	18.410	167
2015	19.271	0
Total	90.782	7.515

Tabla 5-1: Cantidad de Accidentes Localizados en la base de datos SIEC-2 procesada por CONASET.
Fuente: Elaboración propia en base a SIEC-2.

Para describir la información con la que se cuenta, en la Tabla 5-2 se presenta un resumen con los campos de la base de datos y una breve descripción de cada campo.

Campo	Descripción
Año	Año de ocurrencia del accidente
IdAccident	Identificador del accidente
ubicado?	1 si el accidente no fue localizado, 0 si
Fecha	Fecha del siniestro
Comuna	Comuna de ocurrencia del siniestro
Tipo_Accid	Glosa descriptiva del tipo de accidente, se consideran 24 tipos
Tipo_CONASET	Atropello, caída, choque, colisión, incendio u otro tipo. La diferencia entre choque y colisión es que en la primera es un vehículo siniestrado con un objeto, y en la segunda involucra el impacto de 2 o más vehículos de todo tipo
atropello	1 si es accidente atropello, 0 si no
Zona	Ámbito urbano o rural
Ubicación_relativa	Acceso no habilitado, acera o berma, cruce con semáforo apagado, cruce con semáforo funcionando, cruce con señal ceda el paso, cruce con señal pare, cruce regulado por carabinero, cruce sin señalización, enlace a desnivel, enlace a nivel, otros no considerados, plaza de peaje, puente, rotonda, tramo de vía curva horizontal, tramo de vía curva vertical, tramo de vía recta y túnel
Causa_CONASET	Alcohol en conductor, alcohol en pasajero, alcohol en peatón, causas no determinadas, deficiencias viales, desobediencia a señalización, drogas y/o fatiga en conductor, fallas mecánicas, imprudencia del conductor, imprudencia del pasajero, imprudencia del peatón, otras causas, pérdida control vehículo y velocidad imprudente
Causa_Accid	Glosa de causa probable del accidente, se consideran 60 posibles
Calle_Uno	Nombre de la primera calle de referencia
Calle_Dos	Nombre de la segunda calle de referencia
Numero	Numeración frente a la que ocurrió el accidente
Ruta	Nombre de la ruta de referencia donde ocurrió el accidente
Calzada	Bidireccional, Bidireccional con Bandejón o Unidireccional
Fallecidos	Cantidad de fallecidos en el siniestro
Graves	Cantidad de víctimas en condición grave
Menos_Grav	Cantidad de víctimas en condición menos grave
Leves	Cantidad de víctimas en condición leve
Lesionados	Cantidad total de víctimas con algún grado de lesión en el accidente
geomx	Coordenada x de la ubicación del accidente. EPSG:32719, WGS 84 / UTM zona 19S
geomy	Coordenada y de la ubicación del accidente. EPSG:32719, WGS 84 / UTM zona 19S

Tabla 5-2: Campos de la base de datos de accidentes utilizada, derivada de la base de datos SIEC-2.
Fuente: Elaboración propia en base a SIEC-2.

Finalmente se contó con 90.782 registros de accidentes entre los años 2011 y 2015, en los que no siempre se tienen todos los datos de los campos, especialmente para el año 2015, en donde la información con la que se cuenta está incompleta. La información importante está contenida en la ubicación de cada accidente en los campos “geomx” y “geomy”, que se calcularon según los archivos Shapefile entregados por CONASET.

Comuna	Cantidad total de accidentes, periodo 2011-2015
SANTIAGO	10.602
PUENTE ALTO	6.953
PROVIDENCIA	5.426
ÑUÑO A	5.039
MAIPU	4.885
LAS CONDES	4.833
LA FLORIDA	4.348
SAN BERNARDO	3.256
QUINTA NORMAL	3.009
VITACURA	2.922
LA CISTERNA	2.770
PUDAHUEL	2.452
PEÑALOEN	2.412
LA REINA	2.159
MACUL	1.945
LA PINTANA	1.934
SAN JOAQUIN	1.913
ESTACION CENTRAL	1.879
LA GRANJA	1.851
QUILICURA	1.803
CONCHALI	1.795
RECOLETA	1.779
SAN RAMON	1.691
PEDRO AGUIRRE CERDA	1.600
CERRO NAVIA	1.477
HUECHURABA	1.386
EL BOSQUE	1.327
RENCA	1.277
INDEPENDENCIA	1.241
LO PRADO	1.147
LO BARNECHEA	1.091
SAN MIGUEL	1.080
LO ESPEJO	871
CERRILLOS	629
TOTAL	90.782

Tabla 5-3: Total de accidentes para cada comuna.

Fuente: Elaboración propia en base a SIEC-2.

Se destaca de la Tabla 5-3 que las cuatro comunas con mayor cantidad de accidentes, en orden decreciente son Santiago, Puente Alto, Providencia y Ñuñoa, donde Santiago supera en más de 3.600 accidentes a Puente Alto.

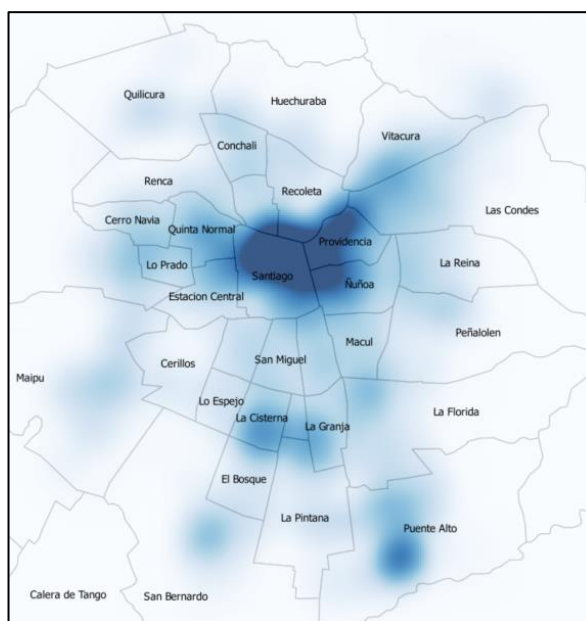


Figura 1: Mapa de calor de los accidentes de tránsito de Santiago.
Fuente: Elaboración propia en base a SIEC-2.

La Figura 1 representa un mapa de calor de los accidentes de tránsito entre los años 2011 y 2015, inclusive. Esta es una representación gráfica en la que se observa que, a mayor densidad de accidentes viales, más oscuro será el color. Este tipo de ilustración se puede implementar desde el software QGIS mediante el uso de la herramienta Estilo, función mapa de calor.

Se puede observar de la Figura 1 que los accidentes tienen mayor presencia en las comunas indicadas anteriormente según la Tabla 5-3, en particular en el sector de la plaza de Puente Alto, el paradero 25 de la Gran Avenida en La Cisterna, y todos los sectores del eje Providencia, Santiago y Ñuñoa.

En el anexo 10.2 se presentan las distribuciones de los accidentes por año de estudio.

Sobre los tipos de siniestros, se observa que hay 55.824 colisiones (61,5%), 17.854 choques (19,7%), 11.019 atropellos (12,1%), 3.145 caídas (3,5%), 1.887 de otro tipo (2,1%), 1.042 volcaduras (1,1%) y solo 11 incendios (0,01%), en total en los 5 años.

La cantidad de siniestros por causas de ocurrencia se pueden observar en la Tabla 5-4 para todos los años, excepto 2015, debido a que no se cuenta con dicha información. Cabe detenerse en un hecho antes mencionado: el reporte de Carabineros solo considera una causa probable, la cual se puede ver fuertemente influenciada según la algidez del tema en la sociedad o simplemente por el criterio del funcionario que registre la información. La experiencia mundial en este tema sugiere reportar más de una, tal cual se comentó al comienzo de esta sección.

Grupo de Causa	Cantidad de Accidentes
Imprudencia Del Conductor	35.137
Otras Causas	9.992
Causas No Determinadas	8.165
Desobediencia A Señalización	7.578
Imprudencia Del Peatón	2.822
Perdida Control Vehículo	2.598
Alcohol En Conductor	2.545
Velocidad Imprudente	1.076
Fallas Mecánicas	567
Imprudencia Del Pasajero	376
Deficiencias Viales	230
Drogas Y/O Fatiga En Conductor	206
Alcohol En Peatón	197
Alcohol En Pasajero	22
TOTAL	71.511

Tabla 5-4: Cantidad de Accidentes por Grupo de causa para todos los años.
Fuente: Elaboración propia en base a SIEC-2.

Sobre la gravedad de los accidentes presentada en la Tabla 5-5, las víctimas resultan con lesiones mayoritariamente leves y un porcentaje menor de consecuencias fatales (2,1% del total de lesionados).

Año	Fallecidos	Graves	Menos Graves	Leves	Lesionados
2011	224	1.378	738	8.624	10.964
2012	210	1.362	644	8.410	10.626
2013	252	1.586	719	8.618	11.175
2014	247	1.570	665	8.247	10.729
2015	234	1.606	631	8.523	10.994
TOTAL	1.167	7.502	3.397	42.422	54.488

Tabla 5-5: Cantidad de víctimas según gravedad de las lesiones, por año.
Fuente: Elaboración propia en base a SIEC-2.

En la Tabla 5-6 se puede observar que la mayoría de las víctimas fatales fueron producidas por accidentes del tipo atropello. Las víctimas graves fueron mayoritariamente producidas por colisiones y atropellos, al igual que las víctimas menos graves. Las víctimas leves fueron mayoritariamente de colisiones, sin despreciar los atropellos y los choques. Esta información corresponde a los años 2011, 2012, 2013 y 2014 por falta de datos en el año 2015. Entre paréntesis se presenta la información de la gravedad de los siniestrados según el total del tipo de accidente.

Tipo de accidente	Fallecidos	Graves	Menos Graves	Leves	Total de Lesionados
Atropello	462 (4,9%)	2.164 (23,0%)	808 (8,6%)	5.991 (63,6%)	9.425 (100%)
Caída	10 (0,4%)	312 (13,0%)	152 (6,3%)	1.934 (80,3%)	2.408 (100%)
Choque	135 (2,4%)	575 (10,0%)	291 (5,1%)	4.729 (82,5%)	5.730 (100%)
Colisión	297 (1,2%)	2.617 (10,7%)	1.387 (5,7%)	20.111 (82,4%)	24.412 (100%)
Incendio	-	-	-	2 (100%)	2 (100%)
Otro Tipo	3 (0,4%)	87 (12,4%)	52 (7,4%)	560 (79,8%)	702 (100%)
Volcadura	26 (3,2%)	141 (17,3%)	76 (9,3%)	572 (70,2%)	815 (100%)
TOTAL	933 (2,1%)	5.896 (13,6%)	2.766 (6,4%)	33.899 (77,9%)	43.494 (100%)

Tabla 5-6: Cantidad de víctimas según gravedad de las lesiones por agrupación del tipo de accidente.
Fuente: Elaboración propia en base a SIEC-2.

Dos de los indicadores de la seguridad vial más utilizados internacionalmente son el riesgo de salud, que se define como la tasa de fallecidos por cada 100.000 habitantes, y el riesgo de tráfico que se define como la cantidad de fallecidos por cada 100.000 vehículos (Nilsson, 2004). La Tabla 5-7 muestra estos valores para los distintos años en estudio para la región metropolitana.

	2011	2012	2013	2014	2015
Fallecidos	360	372	420	431	413
Habitantes RM	6.971.899	7.057.491	7.142.893	7.228.581	7.314.176
Parque vehicular	1.490.951	1.597.762	1.695.442	1.828.033	1.877.667
Riesgo de salud	5,16	5,27	5,88	5,96	5,65
Riesgo de Tráfico	24,15	23,28	24,77	23,58	22,00

Tabla 5-7: Riesgo de salud y de tráfico para la región metropolitana por cada año del estudio.
Fuente: Elaboración propia en base a datos del INE y SIEC-2

Se pueden obtener los accidentes contenidos en las diversas zonas EOD o comunas de Santiago mediante el recurso de herramientas de análisis como las matrices de distancia entre objetos, la unión y diferencia geométrica entre polígonos, intersección de información mediante su georreferenciación, entre otras.

La cantidad de accidentes por zona se muestra en la Figura 2.

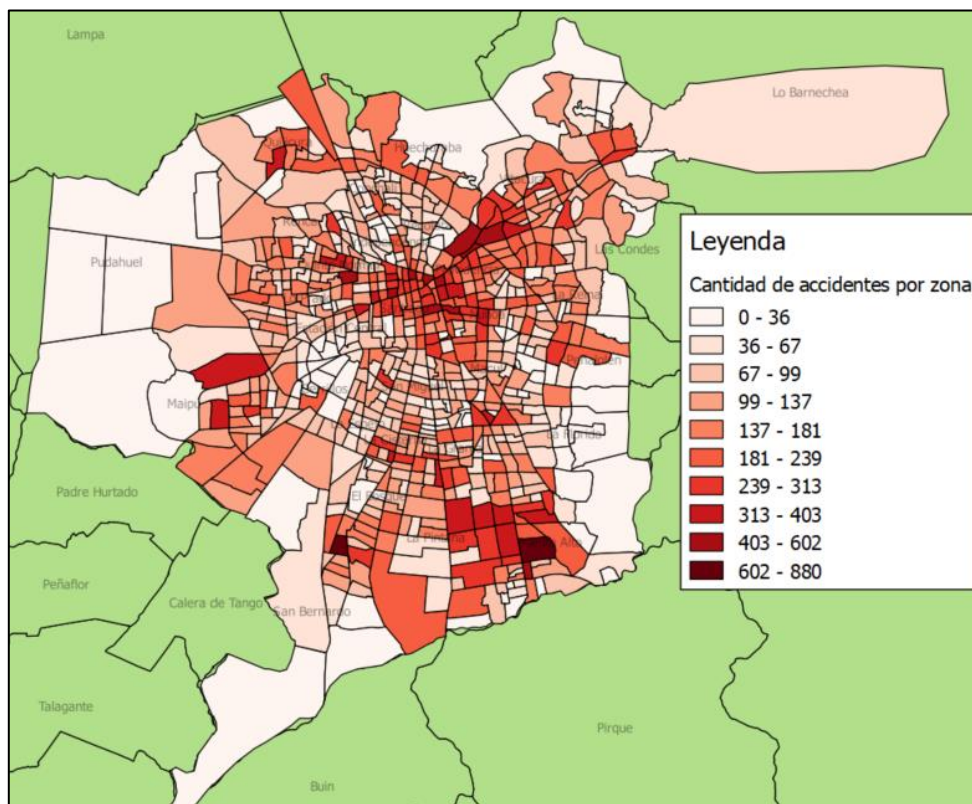


Figura 2: Cantidad de accidentes por zona EOD 2012.
Fuente: Elaboración propia.

5.2 Encuesta Origen-Destino del año 2012 para Santiago

La encuesta origen-destino (EOD) del año 2012 es una base de datos que contiene el patrón de viajes de los habitantes de Santiago. Es el material sobre el cual se desarrollan diversos tipos de políticas públicas relativas al transporte, medio ambiente y desarrollo urbano.

La encuesta y sus resultados se componen del desarrollo de los siguientes componentes (SECTRA, 2014):

1. Encuesta origen-destino de viajes en hogares: consiste en encuestar a 18.000 hogares de Santiago, seleccionados aleatoriamente, en el periodo comprendido entre julio de 2012 y noviembre de 2013.
2. Encuesta de interceptación en cordón externo: esta encuesta complementa la información de la encuesta a hogares con la finalidad de estimar los viajes externos que utilizan el sistema de transporte de Santiago.
3. Catastro de servicios de transporte público: complementa la información asociada al comportamiento de los viajes. Se elaboran catastros de servicios de buses rurales y taxi colectivos.
4. Medición de niveles de servicio de transporte público y privado: mediciones que permitieron determinar los tiempos de viaje de los distintos modos de transporte.
5. Recopilación de información para dos temporadas (normal y estival), diferenciándose en día laboral, sábado y domingo.

El área de estudio contempla 45 comunas del Gran Santiago, en donde se estima que habitan 6,5 millones de personas, con una cantidad estimada de 1.160.000 vehículos particulares, 6.300 buses urbanos, 27.000 taxis básicos, 11.000 taxis colectivos urbanos y 5 líneas de metro con 104 km de vías.

Los principales resultados publicados en el informe ejecutivo del estudio (SECTRA, 2014) son:

1. Cada habitante realiza, en promedio, 2,78 viajes diarios.
2. Cada hogar, en promedio, genera 9 viajes diarios.
3. 29,1% de los viajes son realizados en transporte público.
4. 28% de los viajes son realizados en transporte privado.
5. 34,5% de los viajes son caminando, 25,7% en automóvil y 25% en Transantiago.
6. De ese último 25%, 52% son modo solo bus, 26% combinación bus-metro y 22% solo metro.
7. Los viajes en bicicleta representan el 4% en un día laboral normal.

Uno de los procesos más relevantes de la encuesta es considerar una zonificación acorde a la situación de la ciudad. Estas zonas que se generan son utilizadas en la modelación de accidentes de tránsito. Dado que se consideran solo 34 comunas del Gran Santiago, y la encuesta origen destino abarca más de estas, es necesario considerar solo la zonificación acorde a las comunas de análisis de esta memoria, lo que se muestra en verde en la Figura 3.

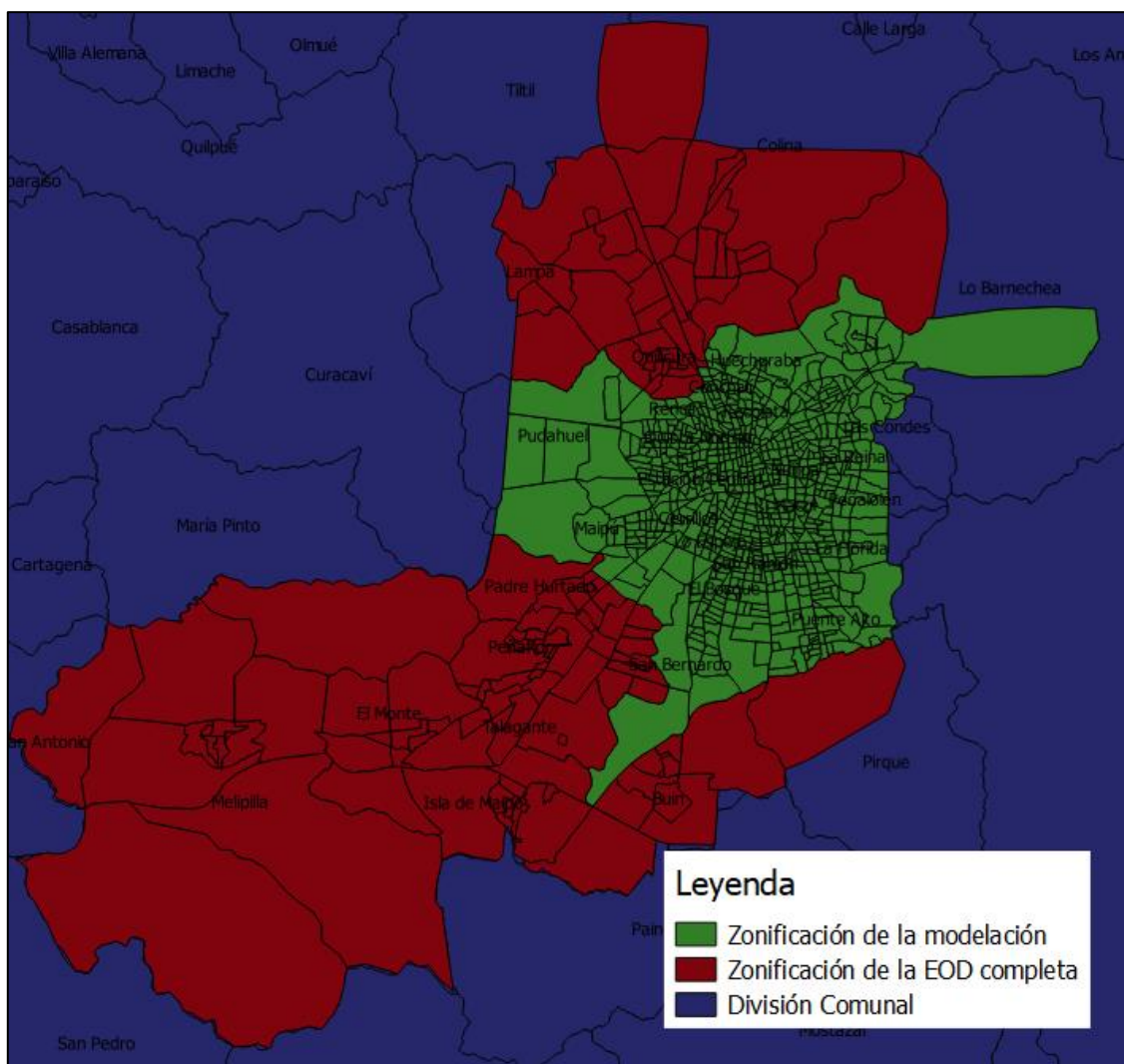


Figura 3: Zonificación considerada en la modelación y zonificación de la Encuesta Origen Destino 2012.
 Fuente: EOD 2012, Biblioteca del Congreso Nacional y elaboración propia.

La base de datos contiene, entre otras tablas, información de cada etapa de un viaje, donde este último se considera como el desplazamiento de una persona desde un origen a un destino, en un cierto modo de transporte, en determinado momento del día. Una etapa del viaje corresponde al desplazamiento en un solo modo, de una proporción de un viaje. Luego, un viaje está constituido por varias etapas en varios modos. Por ejemplo: una persona toma un bus en el paradero más cercano, luego aborda el servicio de metro y termina abordando otro servicio troncal hacia su destino. En este caso se consideran 3 etapas.

Cabe destacar también que la EOD incluye códigos de reporte de las rutas de los servicios del Transantiago que los usuarios declararon que utilizaron en sus viajes. Estos códigos no necesariamente se correlacionan con lo reportado en los programas de operación del Transantiago, explicados más adelante.

Además, la EOD reporta las coordenadas de inicio y fin de un viaje, así como el modo “enteramente caminando”, bicicleta y moto. Con esta información se pueden generar

rutas por donde las personas caminaron y luego generar una estimación del número de peatones en cada zona. Esto mismo se puede realizar para los modos bicicleta y moto.

La cantidad de viajes generados en un día laboral normal, según constata la EOD, son de aproximadamente 18.461.134. Estos viajes generan 20.981.889 etapas, es decir, se generan 2.520.755 transbordos entre todos los modos.

La Tabla 5-8 muestra la cantidad de viajes generados en un día laboral normal por los distintos modos agregados. Se puede observar que los principales modos en la ciudad son la caminata, el auto y los viajes en Transantiago.

Modo Agregado	Cantidad de Viajes	% del total
Caminata	6.363.320	34,5%
Auto	4.748.580	25,7%
Bus TS	2.290.011	12,4%
Bus TS - Metro	1.110.152	6,0%
Otros	826.628	4,5%
Metro	771.149	4,2%
Bicicleta	747.123	4,0%
Taxi Colectivo	530.363	2,9%
Taxi	314.977	1,7%
Bus no TS	308.774	1,7%
Otros - Bus TS	87.856	0,5%
Taxi Colectivo - Metro	81.829	0,4%
Bus no TS - Metro	80.598	0,4%
Otros - Bus TS - Metro	65.485	0,4%
Auto - Metro	47.468	0,3%
Bus TS - Bus no TS	42.510	0,2%
Otros - Metro	29.585	0,2%
Taxi - Metro	14.726	0,1%

Tabla 5-8: Cantidad de viajes generados en un día laboral normal según modo agregado.
Fuente: EOD 2012 y elaboración propia.

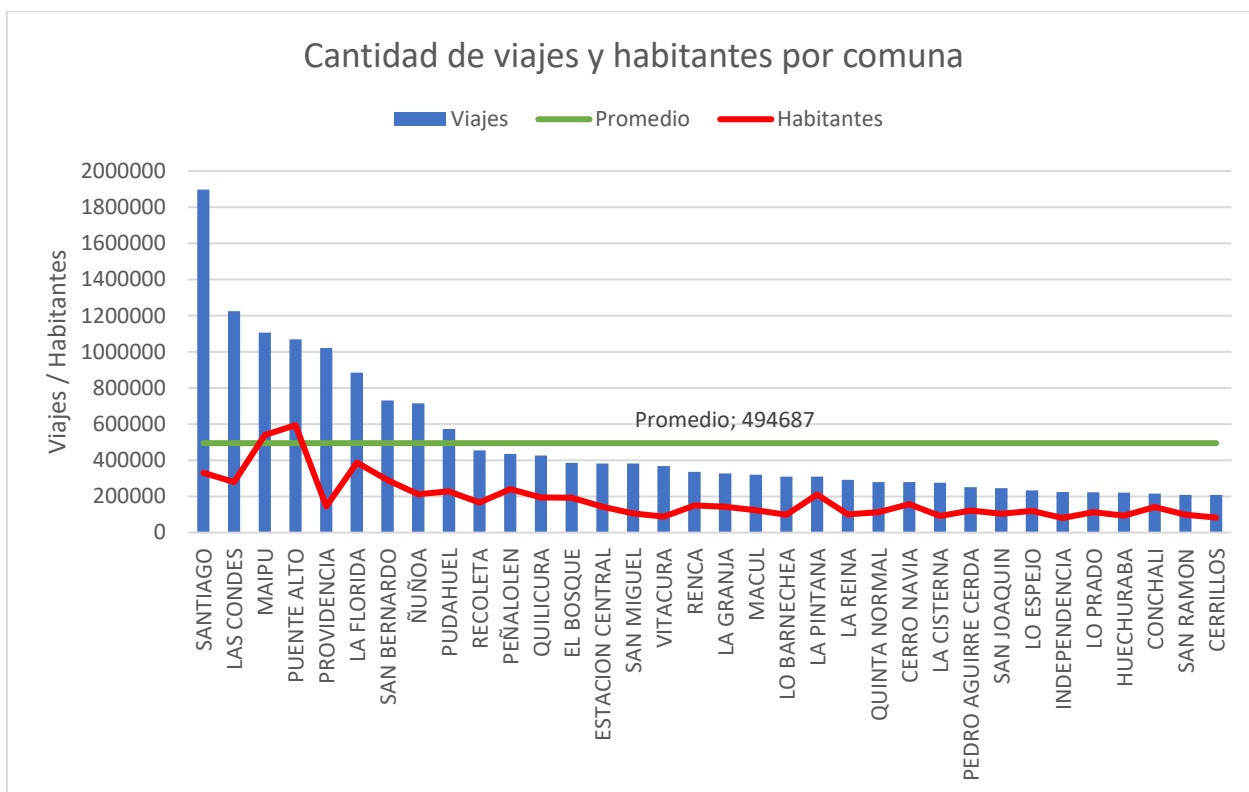


Gráfico 1: Gráfico de cantidad de viajes generados por comuna, incluyendo la cantidad de habitantes.
Fuente: Elaboración propia en base a EOD 2012 y datos demográficos del INE.

El Gráfico 1 muestra que la cantidad de viajes generados en la comuna de Santiago es altísima comparado con el promedio de viajes por comuna, al igual que para Las Condes, Maipú, Puente Alto y Providencia. Solo en estas 5 comunas se generan el 37,6% de todos los viajes generados en las 34 comunas del Gran Santiago. Si extendemos el análisis a las comunas por sobre el promedio, obtenemos que 9 comunas de 34 (26,41%) generan el 54,8% del total de viajes en la ciudad.

De la Figura 4 se puede observar que la distribución de hogares encuestados está más concentrada cerca de las comunas de Santiago, Providencia, Ñuñoa y Puente Alto.

De la Figura 5 se destaca que lo más oscuro del mapa representa un máximo de 7000 viajes en la zona marcada. Esto implica que tanto para los orígenes y los destinos, los viajes se concentran mayoritariamente en los alrededores de Santiago y Providencia.

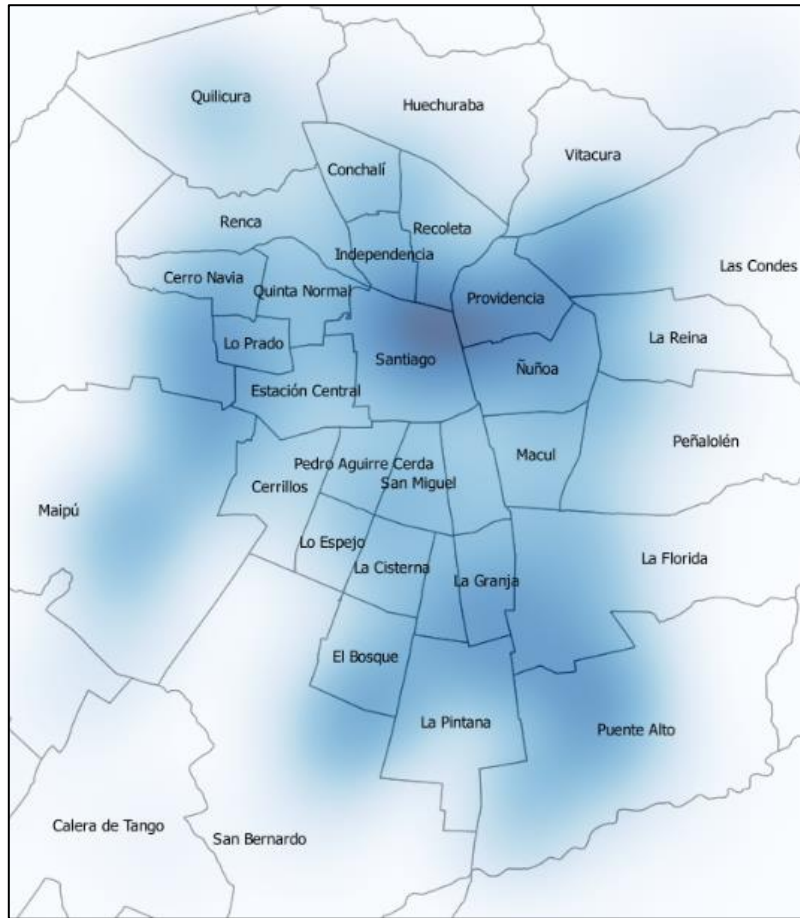


Figura 4: Mapa de calor de los hogares considerados en la EOD 2012.
Fuente: Elaboración propia en base a EOD 2012.

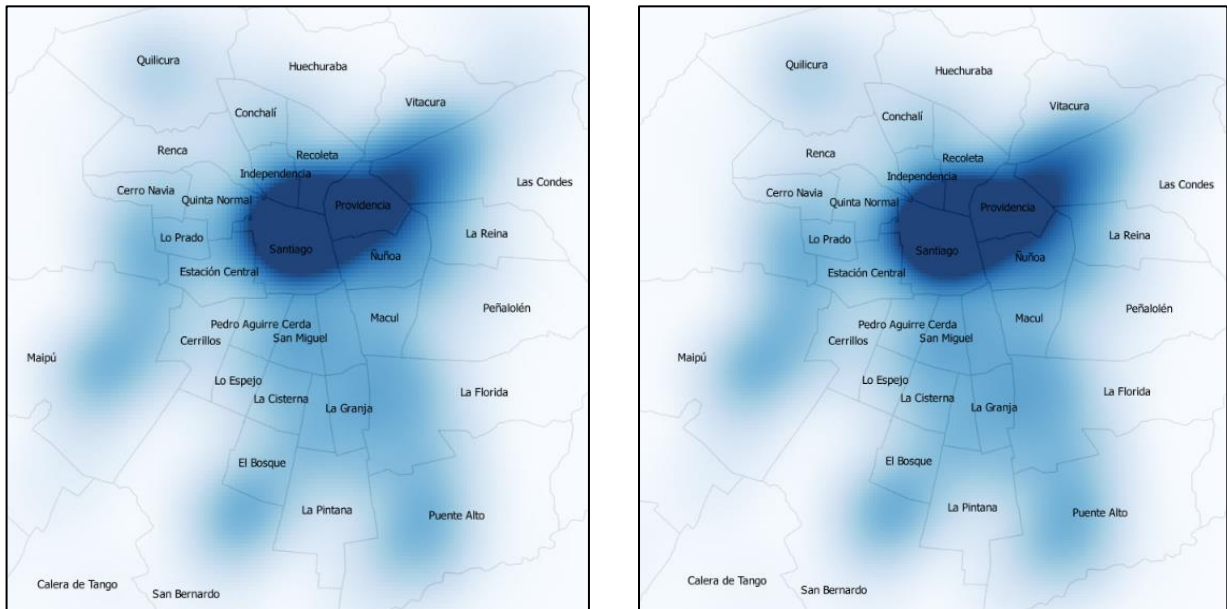
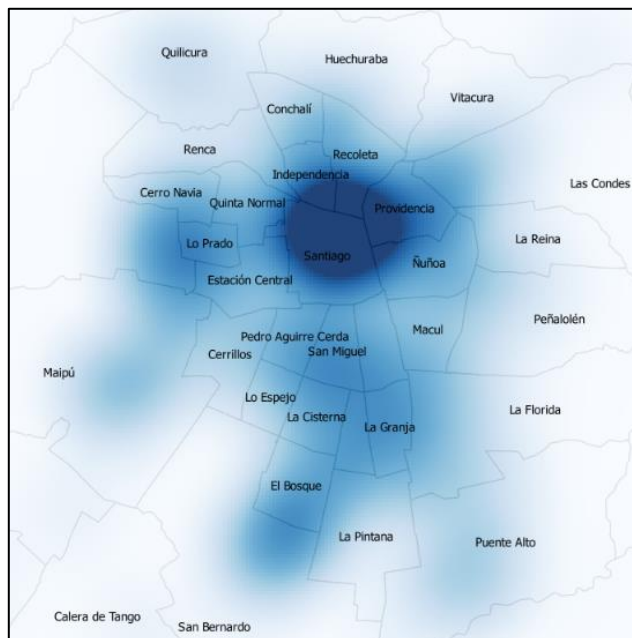
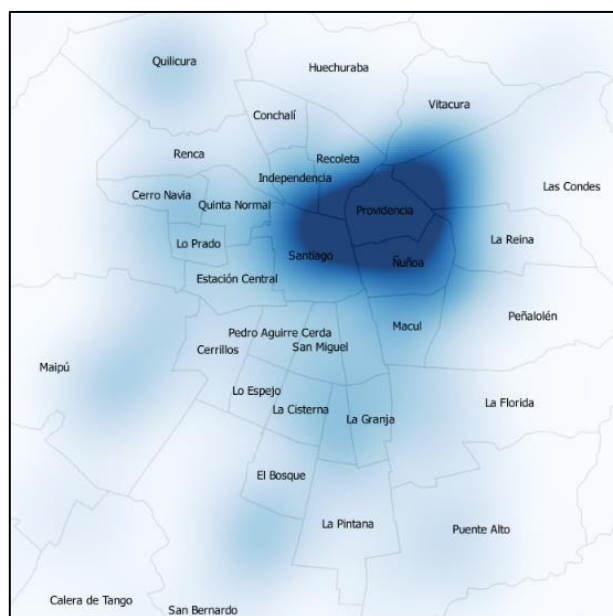


Figura 5: Mapas de calor de los orígenes (izquierda) y destino (derecha) de los viajes.
Fuente: Elaboración propia en base a EOD 2012.

De la Figura 6 y Figura 7 se puede observar el comportamiento de los viajes en modo caminata y bicicleta. Para el caso de caminata, lo más oscuro representa un máximo de 1500 viajes, mientras que, para bicicleta, un máximo de 300. La mayoría de los viajes en caminata son generados en la comuna de Santiago, y los viajes en bicicleta son mayoritariamente producidos en la comuna de Providencia. Esto en términos absolutos sobre el total de los viajes en cada modo.



*Figura 6: Mapa de calor de los viajes generados en caminata.
Fuente: Elaboración propia en base a EOD 2012.*



*Figura 7: Mapa de calor de los viajes en bicicleta.
Fuente: Elaboración propia en base a EOD 2012.*

5.2.1 Procesamiento de los datos

De la información con la que se cuenta, se generan los archivos necesarios para poder generar rutas mínimas entre los distintos pares origen y destino. Esto se realiza para los modos caminata, bicicleta y moto. El objetivo de esto es saber la carga de flujo por modo por zona.

El proceso se basó en el uso del sistema de gestión de bases de datos relacional orientado a objetos llamado Postgres. Este cuenta con un módulo llamado PostGIS que añade funcionalidad espacial a los datos anexando características como tipos de datos espaciales, índices espaciales y funciones acordes. Este, a su vez, cuenta con una extensión llamada Pgrouting que provee funcionalidades de ruteo geoespacial. En particular, se utilizará el algoritmo Dijkstra implementado en esta solución informática.

La metodología consistió en considerar una red de Santiago generada y revisada minuciosamente para que no existieran discordancias al momento de utilizar el algoritmo. Esto es básicamente que no existan bucles, que no haya calles sin conexiones, que los nodos estén bien constituidos y que la red sea concordante con la realidad. Esta red es la publicada en una tesis de postgrado de la Universidad de Chile, sobre evaluación en red de proyectos de reconstrucción de establecimientos educacionales (Maluenda Nazar, 2012).

Al generar la base de datos espacial con dicha red, se generan automáticamente nodos con coordenadas espaciales. A estos se les asignan los distintos pares origen destino según cercanía y pertenencia a la zona que se indica en la base de datos de la EOD a modo de filtro, pues un punto puede estar más cerca de un nodo, pero no necesariamente puede ser que este se ubique en la zona en que se declara el origen o destino.

Con dichos datos ya cargados en Postgres, se procede a ejecutar el algoritmo para los distintos pares origen destino. El proceso detallado, junto con los códigos de consulta, se presentan en el anexo 10.4.

Cabe destacar que en la EOD existían muchos datos de viajes que no poseían coordenadas y/o no tenían declarada la zona de origen o eran de valor 0 o 999, irreconocibles por el diccionario de zonas. A continuación, se presenta una tabla con los problemas y resultados del proceso generado.

	Caminata	Bicicleta	Moto
TOTAL ORIGINAL DE LA BASE DE DATOS EOD 2012	33.827	3.893	547
Sin zona origen (zona declarada de valor 0 o 999)	978	133	17
Sin zona destino (zona declarada de valor 0 o 999)	690	99	16
Sin coordenadas origen	13.381	252	15
Sin coordenadas destino	13.091	217	14
Zona origen 999	6	7	2
Zona destino 999	8	8	2
Zonas origen y destino 0	58	0	0
Zonas con origen cero, pero destino distinto de cero (Sin casos con zona destino 999)	914	123	15
Zonas con destino cero, pero origen distinto de cero	624	88	14
Zonas sin origen ni destino	12.408	126	0
Pares origen destino válidos para rutas mínimas	19.822	3.553	518
Etapas útiles para ser consideradas en la carga por modo por zona	33.230	3.679	518
Porcentaje de pares válidos para rutas mínimas al respecto del total original	95,3%	94,5%	94,7%
Pares OD que generaron rutas mínimas	31.947	3.561	513
PORCENTAJE DE PARES VÁLIDOS PARA LA MODELACIÓN	94,4%	91,2%	93,4%

Tabla 5-9: Detalle de problemas y resultados de las rutas mínimas para caminata, bicicleta y moto.

Fuente: Elaboración propia.

Las etapas que tienen declarada zona origen o destino cero no tienen coordenadas para ser ubicados los puntos, por lo que son descartados. Además, las etapas que tienen declarada zona 999 en el origen y en el destino tienen coordenadas, por lo que no son eliminados, pero se tuvo presente en el análisis posterior.

Existían muchos pares OD sin zona origen ni destino, tal cual se explicita en la Tabla 5-9. Estos pares eran todos viajes intrazonales, por lo que no son considerados en el proceso de rutas mínimas, pero si son agregados como carga para cada modo en donde se ubican. Había también pares origen destino que tenían coordenadas distintas en el origen y destino, pero que también eran intrazonales. Se consideraron en el resultado para la modelación.

Para los viajes que no reportan coordenadas se considera que el centroide de la zona es su coordenada, de manera de incluir estos viajes a pesar de la inexistencia de dicha información.

Finalmente, se generó un factor de ampliación de las rutas mínimas para considerar las etapas que no tenían información en la zona origen o zona destino (zona 999) que se construye como el porcentaje de etapas por modo que no son consideradas, dividido por el total.

Comunas	Carga caminata	Carga Bicicleta	Carga moto
Santiago	882.944	255.979	133.556
Maipú	379.013	97.584	63.731
Providencia	347.258	236.883	55.109
Puente Alto	332.698	95.994	17.319
La Florida	298.524	65.577	41.364
Pudahuel	282.255	33.247	17.414
San Bernardo	277.074	64.884	23.123
Recoleta	267.270	46.808	12.131
Las Condes	261.520	144.594	58.963
Peñalolén	219.167	48.131	32.885
Ñuñoa	218.369	208.485	41.865
El Bosque	177.357	23.797	35.473
Estación Central	167.504	51.188	23.911
Lo Prado	158.639	17.918	2.100
La Granja	157.218	27.592	2.706
Pedro Aguirre Cerda	150.745	24.135	19.045
San Miguel	144.254	33.240	31.102
Cerro Navía	140.100	26.413	3.962
San Joaquín	139.502	51.598	30.368
Independencia	136.551	32.796	18.320
Conchalí	134.929	27.987	10.004
Macul	128.566	55.826	43.526
Cerrillos	119.684	34.793	16.623
Quinta Normal	111.562	38.479	5.198
La Pintana	102.460	23.880	5.424
La Cisterna	101.786	53.913	37.029
Lo Espejo	98.146	17.222	14.782
Renca	97.680	34.015	19.792
Vitacura	97.636	56.826	15.324
San Ramón	83.913	29.814	1.091
Huechuraba	79.908	28.019	2.972
La Reina	70.859	51.838	20.972
Lo Barnechea	47.643	11.322	5.114
TOTAL	6.412.734	2.050.778	862.297

*Tabla 5-10: Carga de caminata, bicicleta y peatones por comuna, promedio por día.
Fuente: Elaboración propia.*

Los datos mostrados en la Tabla 5-10 es la suma de la carga estimada por modo por día promedio de las zonas que pertenecen a cada comuna de estudio. La carga representa la suma de viajes que pasaron por una zona en su ruta, ponderada para un día laboral, sábado o domingo, según si es un día normal o un día estival. Es decir, es una medida de flujo por día para todos los tipos de días según modo por cada zona. Cabe destacar que un viaje comenzando en una zona determinada, puede aportar más de una vez en varias zonas debido a que en su ruta pasó por varias.

Todos los datos analizados están amplificados por factores de expansión reportados en la EOD 2012, los cuales son ajustes de representación de cada viaje según el tipo de persona, el tipo de hogar y el tipo de viaje que se realiza.

5.3 Base de Datos de Bienes Raíces del SII

Una fuente disponible es administrada por el Servicio de Impuestos Internos (SII), llamada Base de Datos de Bienes Raíces (BBRR). Esta consta básicamente de datos sobre los predios de la ciudad, como por ejemplo avalúo fiscal, principal actividad del predio, superficie total, material de construcción, calidad línea de construcción, entre otras. Se realizó una solicitud a través de la ley de transparencia (MINSEGPRES, 2008) al SII para obtener:

1. Catastro físico de los inmuebles de la Región Metropolitana (base de datos catastral de predios rurales y no rurales) para los años desde el 2008 al 2014 por mes y/o semestre según disponibilidad (producto 4 de los Disponibles a Entregar en web SII) con todos los ítems disponibles (comuna, dirección, manzana, predio, destino, superficie de terreno, etcétera).
2. Planos (Shapefile, CAD u otro disponible) de predios y manzanas de las comunas que componen la Región Metropolitana.

La respuesta obtenida fue satisfactoria en el sentido que los datos se recibieron conforme, mas no se cuenta con la información del año 2008. Sobre los planos, se tiene solo la información para 28 comunas de la Región Metropolitana.

Como es información sobre todos los predios de la ciudad, son varios registros por año. En particular, son 20.576.500 roles de avalúo entre los años 2011 y 2014 por semestre, tal cual se detalla en la Tabla 5-11. Esto indica un crecimiento de entre 0,4% y 1,6% entre semestres

Cantidad de roles		
Año	Semestre	de avalúo
2011	1	2.477.660
2011	2	2.518.200
2012	1	2.538.639
2012	2	2.572.116
2013	1	2.582.384
2013	2	2.596.501
2014	1	2.625.156
2014	2	2.665.844

Tabla 5-11: Cantidad de predios por año en cada semestre.
Fuente: Elaboración propia en base a datos SII.

La descripción de la base de datos completa es anexada en el cuerpo 10.3.

La base contiene dos tipos de tablas en las que se muestra la información sobre un rol de avalúo en términos generales y su información sobre el terreno y las construcciones de este. La información relevante para el trabajo de la memoria está contenida en campo 7 sobre el destino principal de uso de suelo. Además, se deben tener en cuenta los números de manzana y predio para luego georreferenciar la información y así ubicar los terrenos en las correspondientes zonas de las bases de modelación.

Sobre el destino principal de los roles de avalúo, se puede notar que el que prepondera es el habitacional, seguido por estacionamiento, bodega y almacenaje, y seguido por comercio. Esta tendencia se repite en todos los años, para todos los semestres, tal cual se puede observar en la Tabla 5-12.

Destino	2011		2012		2013		2014	
	Semestre 1	Semestre 2	Semestre 1	Semestre 2	Semestre 1	Semestre 2	Semestre 1	Semestre 2
COMERCIO	68.186	68.410	68.538	68.765	68.758	68.800	69.122	69.200
DEPORTE Y RECREACIÓN	1.087	1.090	1.099	1.115	1.115	1.121	1.137	1.146
EDUCACIÓN Y CULTURA	5.406	5.451	5.494	5.518	5.513	5.516	5.553	5.560
HOTEL/MOTEL	1.395	1.739	1.897	2.253	2.253	2.258	2.259	2.254
HABITACIONAL	1.809.700	1.835.061	1.845.190	1.863.232	1.873.520	1.882.146	1.896.415	1.920.031
INDUSTRIA	11.215	11.253	11.247	11.265	11.265	11.252	11.262	11.226
BIENES COMUNES	14.649	14.816	14.913	15.065	15.144	15.195	15.327	15.538
BODEGA Y ALMACENAJE	190.907	195.832	198.533	204.895	204.915	207.461	211.619	218.184
MINERÍA	13	13	13	12	12	12	12	12
OFICINA	42.103	43.057	43.701	44.148	44.145	44.193	45.341	46.120
ADMINISTRACIÓN PÚBLICA Y DEFENSA	464	467	467	459	459	459	464	463
CULTO	2.627	2.638	2.653	2.667	2.667	2.676	2.693	2.704
SALUD	1.270	1.278	1.290	1.296	1.296	1.282	1.299	1.297
TRANSPORTE Y TELECOMUNICACIONES	524	538	544	549	548	552	558	557
OTROS NO CONSIDERADOS	3.321	3.411	3.412	3.502	3.501	3.523	3.582	3.609
SITIO ERIAZO	36.814	37.778	37.504	37.763	37.705	37.760	38.440	38.956
ESTACIONAMIENTO	287.979	295.368	302.144	309.612	309.568	312.295	320.073	328.987

Tabla 5-12: Cuantía de roles por cada destino de uso principal, para cada año y semestre
Fuente: Elaboración propia en base a datos SII.

En la Tabla 5-13 se pueden notar la cantidad de metros cuadrados del terreno según el destino de uso principal de cada predio, para cada año. En ella se puede observar que la preponderancia de uso está en el destino habitacional, seguido por sitios eriazos e industria. Por otro lado, el de menor uso es el de hotel/motel y estacionamiento.

Destino	2009		2010		2011		2012	
	Semestre1	Semestre2	Semestre1	Semestre2	Semestre1	Semestre2	Semestre1	Semestre2
COMERCIO	33.180.307	33.463.995	34.214.514	34.323.512	34.311.347	34.491.756	34.810.423	35.132.233
DEPORTE Y RECREACIÓN	29.990.504	29.944.504	29.901.006	29.960.736	29.960.736	30.016.312	30.232.369	30.262.200
EDUCACIÓN Y CULTURA	26.884.019	26.953.188	27.000.595	27.143.998	27.139.141	27.174.673	27.299.040	27.420.992
HOTEL/MOTEL	710.961	729.152	749.672	759.350	759.350	781.047	782.849	784.917
HABITACIONAL	365.920.188	369.751.430	372.933.668	373.852.649	375.094.961	377.146.059	378.387.657	380.351.721
INDUSTRIA	65.040.708	65.016.428	63.973.173	63.995.111	63.995.111	64.129.814	64.375.878	64.485.926
BIENES COMUNES	37.350.808	38.468.817	39.331.408	40.192.123	41.157.965	41.493.406	42.837.824	44.920.313
BODEGA Y ALMACENAJE	13.189.376	14.239.218	14.818.351	15.218.581	15.244.990	15.707.988	15.833.142	15.981.578
MINERÍA	3.258.041	2.912.441	2.912.441	2.731.341	2.731.341	2.731.341	2.731.341	2.731.341
OFICINA	7.018.331	7.338.200	7.378.562	7.672.038	7.670.412	7.803.315	7.739.565	7.831.260
ADMINISTRACIÓN PÚBLICA Y DEFENSA	5.642.141	5.656.668	5.656.367	5.607.149	5.607.149	5.606.189	5.931.653	5.863.044
CULTO	4.224.667	4.232.627	4.269.685	4.293.847	4.293.847	4.304.984	4.347.625	4.387.203
SALUD	4.142.082	4.171.028	4.216.167	4.249.326	4.249.326	4.245.118	4.283.454	4.263.740
TRANSPORTE Y TELECOMUNICACIONES	5.642.656	5.728.126	5.757.357	5.799.003	5.793.311	5.837.196	6.296.425	6.298.537
OTROS NO CONSIDERADOS	58.311.899	59.777.990	59.029.103	59.093.092	59.076.151	59.260.475	58.939.918	58.913.436
SITIO ERIAZO	198.479.725	204.181.184	203.918.963	206.287.752	205.935.423	206.828.297	212.795.666	214.621.470
ESTACIONAMIENTO	1.497.922	1.526.891	1.479.156	1.478.630	1.478.667	1.497.724	1.552.074	1.558.927

Tabla 5-13: Cantidad de metros cuadrados por destino de uso principal del predio para cada año y semestre.

Fuente: Elaboración propia en base a datos SII.

Toda la información anterior no tiene relevancia si no se puede georreferenciar, tal cual se mencionó anteriormente. Este proceso demanda un método de alto costo, por lo que se prosigue a utilizar información existente ya referenciada, que, si bien solo se cuenta para 28 comunas debido a la solicitud por ley de transparencia (Presidencia, 2008), se pudo contar con el resto gracias a Tomás Cox, académico de la Facultad de Arquitectura y Urbanismo de la Universidad de Chile, y a Ricardo Hurtubia, académico de doble afiliación a la Escuela de Arquitectura y al departamento de Ingeniería de Transporte y Logística de la Pontificia Universidad Católica de Chile.

Ellos facilitaron una base de datos georreferenciada de la información de bienes raíces para la mayoría de Santiago. Esta consiste en un consolidado de archivos con datos a nivel de manzana de todas las comunas, excepto Quilicura. Es por esta razón que en la modelación de accidentes se decide omitir la comuna de Quilicura.

En la Figura 8 se muestra un mapa de la información con la que finalmente se cuenta.



*Figura 8: Mapa de la información contenida en el archivo Shapefile de bienes raíces.
Fuente: Elaboración propia en base a datos facilitados por Tomas Cox y Ricardo Hurtubia.*

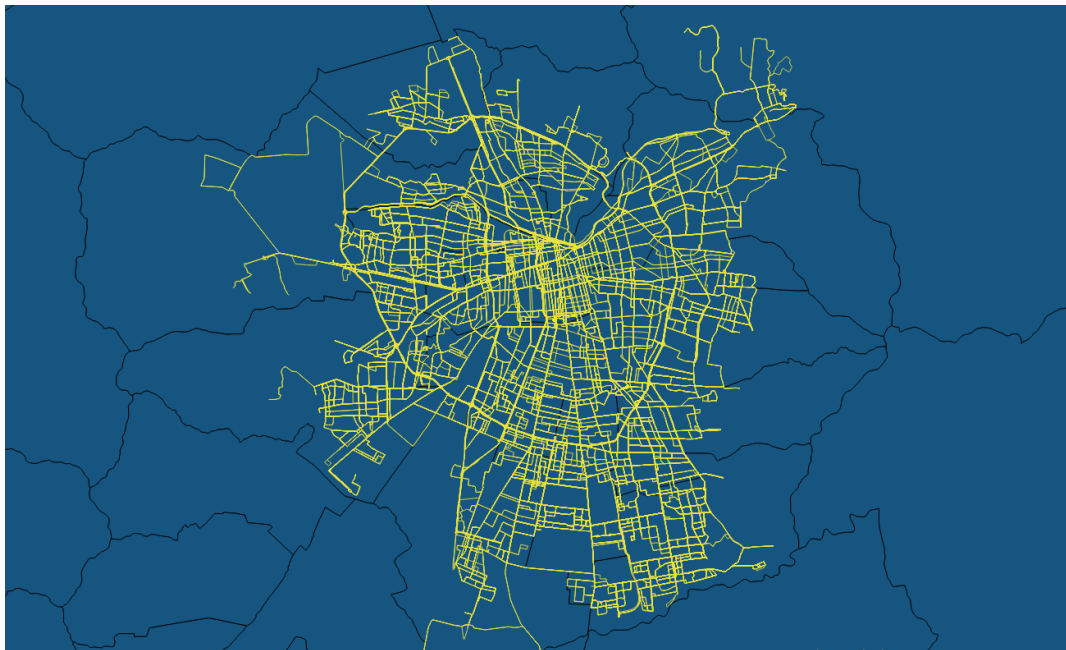
5.4 Programas de Operación del Transantiago

El Directorio de Transporte Metropolitano (DTPM), entidad dependiente del Ministerio de Transportes y Telecomunicaciones, tiene como objetivo articular, coordinar y hacer seguimiento de las acciones, programas y medidas tendientes a gestionar el transporte público de la ciudad de Santiago. En este sentido es que reportan los planes de operación de los servicios de transporte que los concesionarios deben cumplir para el Transantiago. Estas son las guías que definen y regulan las condiciones y características de los servicios, entre ellas: las rutas o trazados de los servicios, los paraderos, las frecuencias, capacidades, horarios de salida, salidas extraordinarias, entre otras.

Los programas de operación son información pública y se encuentran disponibles en la página web del directorio¹, para distintos trimestres desde el año 2011.

Para el desarrollo de esta memoria, se trabaja con los archivos de referencias geográficas para los paraderos y los servicios del segundo semestre del 2012, así como también para el primer y segundo semestre del 2013 y 2015, debido a que son las fechas en las que se reporta que se realiza la encuesta origen destino, como se explica en el apartado siguiente, además de utilizar el año 2015 como parte del análisis de los ajustes de los modelos.

A modo de visualizar la cobertura, se muestra en la Figura 9 las líneas de transporte público para el primer semestre del 2013.



*Figura 9: Cobertura de las líneas de Transporte público, programa de operación primer semestre de 2013.
Fuente: DTPM y elaboración propia.*

¹ <http://www.dtpm.gob.cl/index.php/plan-operacional-historico2>

El desafío de trabajar con estos datos consiste en que en el archivo Shapefile existe información de las rutas desagregadas a nivel de todas las variantes posibles del servicio. Esto debido a que el DTPM lo requiere para la realización del pago por el servicio a los operadores.

Esta información se debe contrastar con la existente en los consolidados de información de frecuencias de los servicios. En estos archivos se cuenta con los códigos de los servicios internos del Transantiago, así como también con los códigos a nivel usuario. Además, contienen el detalle de la frecuencia de ida y retorno para todo un día laboral, sábado y domingo. En los días feriados se ejecutan las mismas frecuencias que en un domingo.

El DTPM cruza ambas informaciones para entregar información consistente con los archivos Shapefile. Sin embargo, en los años de estudio no se reportan estos vínculos, por lo que es necesario realizarlos de manera previa al resumen de información relevante para la modelación.

Como en los archivos Shapefile se cuenta con un código de ruta, además del código del servicio y los nombres de estos, se procede a igualar la información con el consolidado de rutas teniendo en cuenta las siguientes salvedades:

- Los archivos de consolidado de información indican la frecuencia de todo un día, por lo que no es necesario considerar la frecuencia de un servicio nocturno.
- Algunos servicios de inyección (cuando una ruta ya está a máxima capacidad, los operadores inyectan buses al servicio para cubrir la demanda), la variante de ciclovías, vías reversibles y las variantes de ferias comerciales de los servicios no son considerados, debido a que se asume que las variaciones de las rutas son insignificantes a nivel de zona, es decir, aunque un servicio varíe su ruta, no lo hará de manera significativa por otra zona que no sea la del servicio normal.
- Si un servicio no se encuentra en el consolidado de frecuencias, pero si está en el archivo Shapefile, se procede a una búsqueda manual del servicio y a averiguar por si el servicio cambió de nombre. Del proceso anterior se logra hallar varias frecuencias de servicios, pero con otras no se pudo, y por lo tanto no son considerados en las frecuencias de buses.
- Dado que las frecuencias son para un tipo de día particular, al ingresarlos a la base de datos de modelación, se observan la cantidad de días laborales, sábado, domingo y feriados en los años de estudio, y se amplifican para poder obtener una suma ponderada de flujos de buses por día.

De todo el proceso anterior, se obtienen las frecuencias de buses por día promedio de la semana para cada zona de estudio, en donde el promedio de esta información es 1.147,4 buses por zona por día. La frecuencia por zona se muestra en la Figura 10.

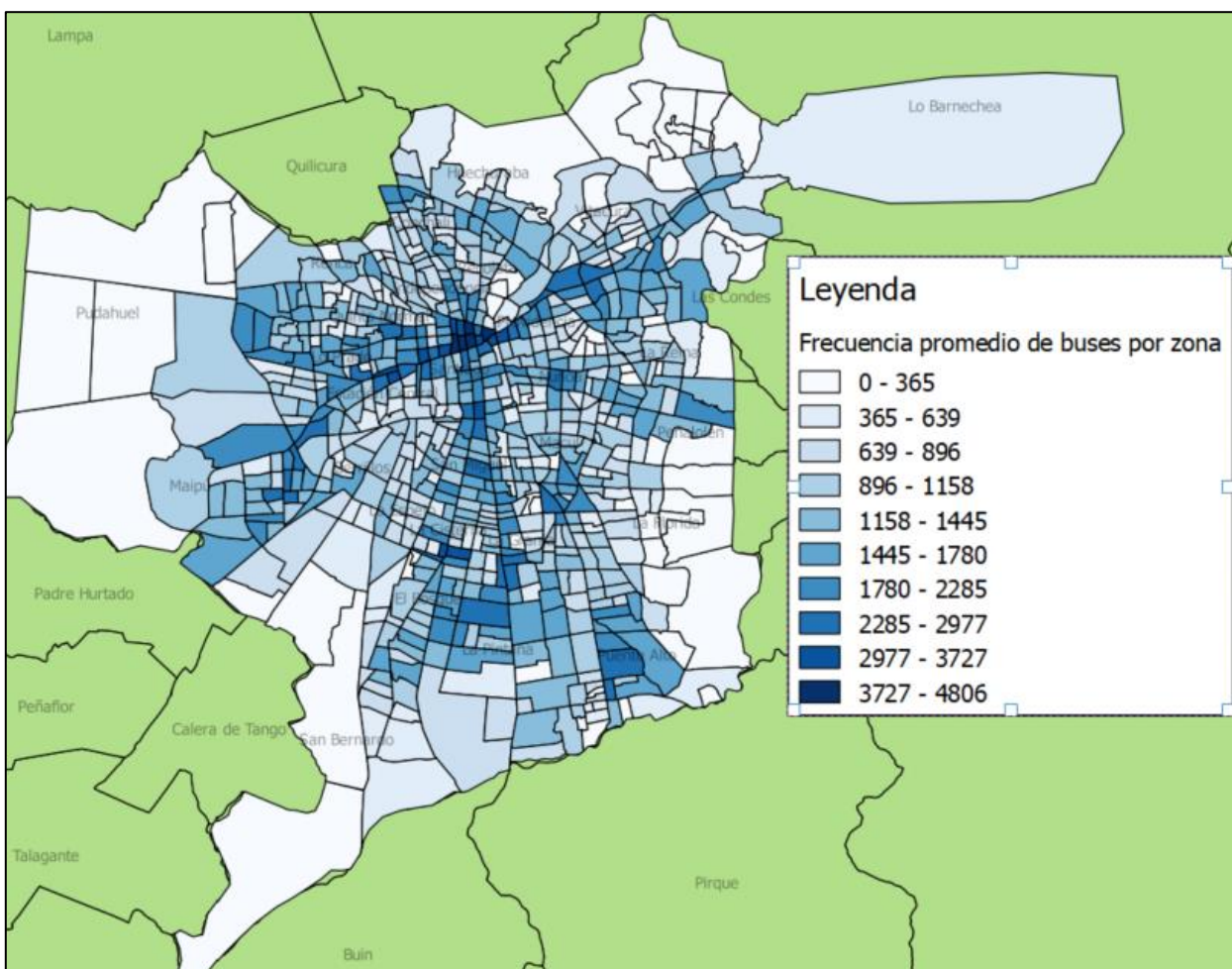


Figura 10: Cantidad de buses promedio por día para cada zona.
Fuente: Elaboración propia en base a datos DTPM.

Parte de las obligaciones que tienen las empresas operadoras del Transantiago consiste en cumplir con los requerimientos del sistema. Por esta razón se generan indicadores de cumplimiento de manera quincenal o mensual según el caso por parte del Centro de Monitoreo de Buses CMB, del DTPM. Esta es la unidad que constituye, desde la perspectiva del diseño, seguimiento, coordinación y planificación, al mejoramiento operacional de los servicios de transporte público, de manera que resuelvan adecuadamente las necesidades de movilidad de sus usuarios.

En particular, se produce el Indicador de Cumplimiento de Frecuencia (ICF) presentado en la Tabla 5-14, el cual mide en términos porcentuales cuál ha sido el nivel de cumplimiento por cada unidad de negocio u operador. Esta información puede ser obtenida desde la plataforma web del DTPM² mediante los informes del ranking de calidad de servicio para los distintos trimestres de los años desde 2012.

² <https://www.dtpm.cl/index.php/2013-04-29-20-33-57/ranking-calidad-de-servicio>

Operadores	Jul-Sep 2012	Oct-Dic 2012	Ene-Mar 2013	Abr-Jun 2013	Jul-Sep 2013	Oct-Dic 2013	Ene-Mar 2015	Abr-Jun 2015	Jul-Sep 2015	Oct-Dic 2015
MetBus	97,9	96,6	99,1	99,4	99,3	99,1	98,9	99,0	98,8	98,8
Vule	95,3	96,6	97,6	97,9	98,0	96,8	96,8	97,6	97,1	96,7
Redbus	95,0	94,4	95,5	96,8	97,5	96,6	95,6	94,6	94,6	95,0
STP	94,8	96,7	96,8	96,8	97,4	95,7	96,1	96,6	96,6	95,5
Subus	93,0	92,3	96,5	96,6	97,0	95,9	94,0	94,4	94,2	93,8
Alsacia	92,8	91,3	93,1	91,7	93,2	92,4	92,2	91,9	91,0	91,7
Express	90,3	91,9	95,2	94,3	96,7	95,3	93,8	88,4	93,0	92,8

Tabla 5-14: Indicadores de Frecuencia ICF para un día completo según operadores del Transantiago en distintos años y trimestres.

Fuente: Elaboración propia en base a informes DTPM.

Con esta información se construyen ICF semestrales para poder amplificar los flujos de buses debido a que se calcula según las frecuencias teóricas que no necesariamente representan la realidad. Esta construcción se muestra en la Tabla 5-15.

Empresa	Unidad de Negocio	2012-2	2013-1	2013-2	2015-1	2015-2
MetBus	5	97,25	99,25	99,22	98,95	98,80
Vule	3	95,95	97,75	97,41	97,20	96,90
Redbus	6	94,70	96,15	97,05	95,10	94,80
STP	7	95,75	96,80	96,54	96,35	96,05
Subus	2	92,65	96,55	96,45	94,20	94,00
Alsacia	1	92,05	92,40	92,81	92,05	91,35
Express	4	91,10	94,75	95,98	91,10	92,90

Tabla 5-15: ICF por semestre según unidad de negocio.

Fuente: Elaboración propia en base a informes DTPM.

Amplificando los flujos de buses por este indicador de frecuencia se obtendrán flujos más cercanos a la realidad. El promedio de variación al respecto de las frecuencias teóricas, por zona de modelación, es de -4,43% con una desviación estándar de 1,05%, una variación mínima de -7,07% y un máximo de -1,42% excluyendo la variación nula.

5.5 Modelo de simulación basada en agentes a gran escala

Para poder predecir cómo reaccionan los usuarios del sistema de transporte frente a diversos estímulos, es necesaria la modelación a nivel usuario considerando las interacciones que ocurren al utilizar la infraestructura vial. En la tesis (Camus, 2017) se realiza una modelación basada en agentes (o usuarios) generando simulaciones que consideran una mejora de la población incluida y ajustes de calibración sobre los flujos y particiones modales observadas. Se genera en base al modelo MATSim (Kickhofer et al ,2016) que es un marco de simulación de transporte basado en actividades diseñado para simular escenarios a gran escala, cuyo caso es la ciudad de Santiago de Chile.

El producto por utilizar del antes mencionado trabajo es una red de arcos geoposicionados del gran Santiago con un nivel simulado de flujo de vehículos particulares para todas las horas del día. Los flujos en el modelo MATSim Santiago fueron calibrados usando conteos de flujo reportados en el informe de medición de aforos realizado por DICTUC (SECTRA, 2013). Esto permite obtener una medida de nivel de carga de flujo de vehículos particulares por día por zona en que se quiera analizar.

El flujo promedio por arco es de aproximadamente 401 vehículos por día y su cobertura se puede observar en la Figura 11.

Cada arco posee información de su largo y de la velocidad a flujo libre. Estos datos se utilizarán para realizar una ponderación de flujos y probar su significancia en los modelos estadísticos de accidentes.

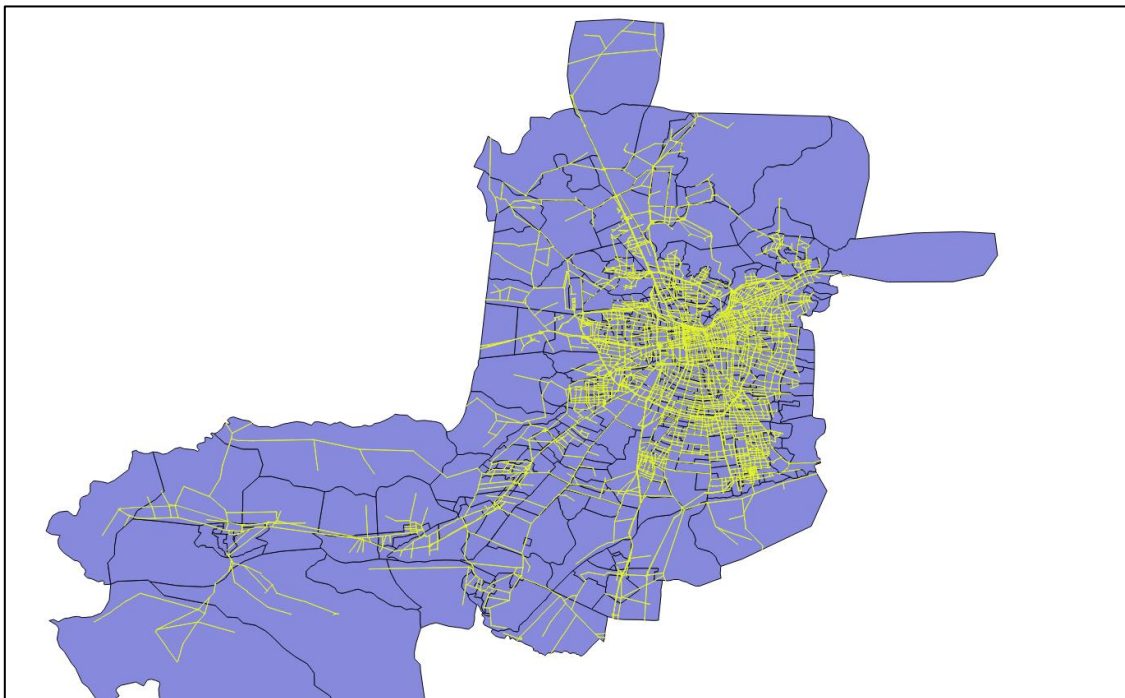


Figura 11: Arcos de flujo vehicular.

Fuente: Elaboración propia en base a tesis de Leonardo Camus.

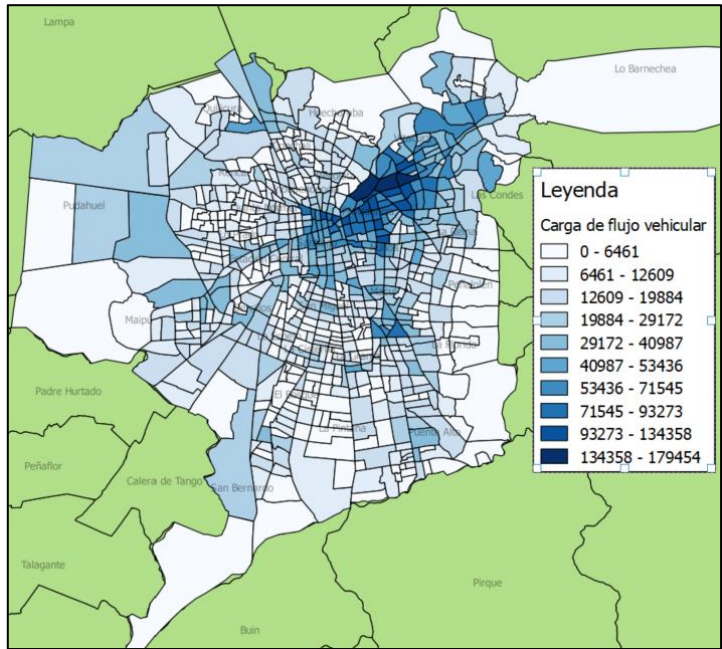


Figura 12: Carga de flujo vehicular.
 Fuente: Elaboración propia en base a tesis de Leonardo Camus

De la Figura 12 se nota que mientras más oscura la zona, mayor flujo. Se puede destacar entonces que las zonas de mayor flujo se encuentran en el centro de Santiago, junto con la comuna de Providencia y Las Condes.

En la Figura 13 se muestra la densidad por arco del flujo vehicular de la red.

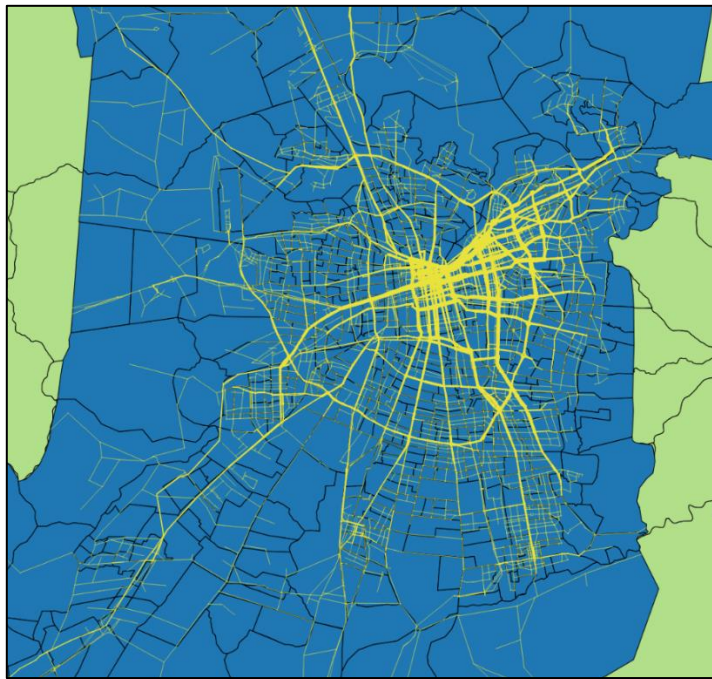


Figura 13: Flujo vehicular con densidad por arco.
 Fuente: Elaboración propia en base a tesis de Leonardo Camus

5.6 Aforos DICTUC

El área de soluciones de movilidad de DICTUC S.A. desarrolló mediciones de flujo y tasas de ocupación de buses, automóviles y otros modos de transporte (SECTRA, 2013). El objetivo principal fue caracterizar los viajes del área urbana de la ciudad de Santiago y de las personas que los realizan. Abarca 45 comunas de la región metropolitana y se realizaron mediciones durante el año 2012.

En este se observan, entre otros cálculos, la tasa de ocupación de los vehículos y buses de la ciudad. Con esto se puede obtener un ajuste de la cantidad de personas que viajan por tipo de vehículo.

Se cuenta con las bases de datos de los resultados de la medición realizada en este estudio. Consta principalmente de mediciones de tipo continuas, periódicas y puntos complementarios. De estos se puede obtener la cantidad de vehículos según su nivel de ocupación y luego calcular, según la flota observada, la tasa de ocupación, como se reportar en la Tabla 5-16.

Categoría	Tipo de vehículo	Ocupación (%)						Cantidad de vehículos observados	Total de pasajeros	Tasa de ocupación	Flota
		1/A	2/B	3/C	4/D	5/E	5+				
C01	Particulares	67%	26%	5%	2%	1%	0%	3.110.074	4.498.913	1,45	-
C02	Taxis	59%	30%	8%	2%	1%	0%	272.431	426.370	1,57	-
C03	Taxis Colectivos	24%	17%	13%	13%	33%	1%	106.567	339.246	3,18	-
C04	Bus Articulado	5%	12%	22%	24%	38%		40.185	1.548.545	38,54	1.441
C05	Bus Grande	8%	12%	20%	21%	38%		71.504	2.239.951	31,33	3.512
C06	Bus Mediano	5%	8%	13%	24%	51%		10.433	231.589,2	22,20	554
C07	Taxi Bus	8%	9%	15%	26%	41%		20.620	432.912,2	20,99	789

Tabla 5-16: Resumen de datos para calcular la tasa de ocupación.
Fuente: Elaboración propia en base a aforos DICTUC (SECTRA, 2013).

La ocupación reportada en las tablas anteriores consta de un número y una letra. El número corresponde a la cantidad de pasajeros por modo, y las letras se corresponden con el nivel de flujo observado en los buses.

Debido a la incapacidad de observar exactamente la cantidad de pasajeros por bus, se considera una agregación que se reajusta en el informe del estudio como una medida cualitativa. Este nivel de ocupación se considera según la Tabla 5-17.

Categoría	Nivel Ocupacional
A	Vehículo completamente ocupado, con pasajeros "casi colgando".
B	Más de la mitad del pasillo con pasajeros de pie.
C	Menos de la mitad del pasillo con pasajeros de pie
D	Más de la mitad de los asientos ocupados
E	Menos de la mitad de los asientos ocupados

Tabla 5-17 :Categorías asociadas a los diferentes niveles de ocupación de buses.
Fuente: (SECTRA, 2013)

Cada una de estas categorías debe asociarse a un valor promedio de tasa ocupacional para cada tipo de bus urbano. La metodología utilizada en el estudio considera medir simultáneamente la tasa de ocupación cualitativa y el número de pasajeros a bordo de

cada vehículo, y con esto actualizar los valores del factor ocupacional para los buses. El resultado de este ajuste se muestra en la Tabla 5-18.

Tipo de bus	Factor ocupacional (pax/bus)				
	A	B	C	D	E
Bus Articulado	139,0	83,1	48,7	30,6	9,5
Bus Grande	88,7	61,8	42,4	24,3	7,5
Bus mediano	67,5	50,3	41,7	24,6	7,5
Taxi bus	51,2	47,4	30,5	21,1	5,4

*Tabla 5-18: Factores de ocupación por tipo de bus agregado.
Fuente: (SECTRA, 2013)*

Con esto se puede obtener la tasa de ocupación por tipo de bus, considerando su respectiva tasa de ocupación según categoría, y luego ponderando por la flota con el fin de obtener un promedio de la tasa de ocupación de los buses de la ciudad.

Finalmente, el resultado estimado es que la tasa de ocupación para los vehículos particulares es de 1,45 y la tasa de ocupación de los buses es de 30,88.

Esta información será luego utilizada en los resultados de los modelos para poder estimar cuál es la participación de un usuario de automóviles y buses en la ocurrencia de siniestros viales.

6 Modelación

6.1 Variables explicadas

Las variables que finalmente fueron recopiladas de las diversas fuentes de información presentadas anteriormente se resumen en la Tabla 6-1.

De dicha tabla, las variables explicadas a utilizar a nivel zonal para el año 2013 son:

- Cantidad de accidentes totales (acc_tot_2013)
- Cantidad de accidentes tipo atropello (acc_at_2013)
- Cantidad de accidentes tipo no atropello (acc_na_2013)
- Cantidad de lesionados (acc_les_2013)
- Cantidad de fallecidos (acc_fall_2013)

La elección del periodo de análisis se realiza considerando que la mayoría de los datos obtenidos son para dicho año.

6.2 Variables explicativas

Las variables explicativas consideran una agregación del tipo dummy para la identificación de comunas, donde cada zona tendrá un valor 1 si es que ella pertenece a cada una de las comunas de análisis.

Por otro lado, se consideran las cantidades de predios por destino de uso de suelo y variables relativas a los flujos en diversos modos, junto con la cantidad de transbordos realizados en cada zona, la suma del avalúo fiscal de todos los predios, el área de una zona, la vialidad presente y una agregación para el flujo vehicular que considera los kilómetros recorridos por cada vehículo utilizada regularmente, llamada vehículo-kilómetro, representando una medida de flujo que intrínsecamente se relaciona también con los kilómetros recorridos.

Las variables relativas al flujo deben ser recalculadas mediante el uso de la función logaritmo debido a la especificación matemática presentada en la Ecuación 13 y Ecuación 14.

Cabe destacar que, para la estimación de modelos, el software utilizado fue SPSS de la empresa IBM en su versión 23. El procedimiento en la herramienta de modelación lineal generalizada se muestra en el anexo 10.7.

Variable	Descripción
DummyStgo	1 si la zona pertenece a la comuna de Santiago, 0 si no
DummyCerrillos	1 si la zona pertenece a la comuna de Cerrillos, 0 si no
DummyCerroNavia	1 si la zona pertenece a la comuna de Cerro Navia, 0 si no
DummyConchali	1 si la zona pertenece a la comuna de Conchalí, 0 si no
DummyElBosque	1 si la zona pertenece a la comuna de El Bosque, 0 si no
DummyEstacionCentral	1 si la zona pertenece a la comuna de Estación Central, 0 si no
DummyHuechuraba	1 si la zona pertenece a la comuna de Huechuraba, 0 si no
DummyIndependencia	1 si la zona pertenece a la comuna de Independencia, 0 si no
DummyLaCisterna	1 si la zona pertenece a la comuna de La Cisterna, 0 si no
DummyLaFlorida	1 si la zona pertenece a la comuna de La Florida, 0 si no
DummyLaGranja	1 si la zona pertenece a la comuna de La Granja, 0 si no
DummyLaPintana	1 si la zona pertenece a la comuna de La Pintana, 0 si no
DummyLaReina	1 si la zona pertenece a la comuna de La Reina, 0 si no
DummyLasCondes	1 si la zona pertenece a la comuna de Las Condes, 0 si no
DummyLoBarnechea	1 si la zona pertenece a la comuna de Lo Barnechea, 0 si no
DummyLoEspejo	1 si la zona pertenece a la comuna de Lo Espejo, 0 si no
DummyLoPrado	1 si la zona pertenece a la comuna de Lo Prado, 0 si no
DummyMacul	1 si la zona pertenece a la comuna de Macul, 0 si no
DummyMaipu	1 si la zona pertenece a la comuna de Maipú, 0 si no
DummyÑuñoa	1 si la zona pertenece a la comuna de Ñuñoa, 0 si no
DummyPAC	1 si la zona pertenece a la comuna de Pedro Aguirre Cerda, 0 si no
DummyPeñalolen	1 si la zona pertenece a la comuna de Peñalólen, 0 si no
DummyProvidencia	1 si la zona pertenece a la comuna de Providencia, 0 si no
DummyPudahuel	1 si la zona pertenece a la comuna de Pudahuel, 0 si no
DummyQtaNormal	1 si la zona pertenece a la comuna de Quinta Normal, 0 si no
DummyRecoleta	1 si la zona pertenece a la comuna de Recoleta, 0 si no
DummyRenca	1 si la zona pertenece a la comuna de Renca, 0 si no
DummySanJoaquin	1 si la zona pertenece a la comuna de San Joaquín, 0 si no
DummySanMiguel	1 si la zona pertenece a la comuna de San Miguel, 0 si no
DummySanRamon	1 si la zona pertenece a la comuna de San Ramón, 0 si no
DummyVitacura	1 si la zona pertenece a la comuna de Vitacura, 0 si no
DummyPuenteAlto	1 si la zona pertenece a la comuna de Puente Alto, 0 si no
DummySanBernardo	1 si la zona pertenece a la comuna de San Bernardo, 0 si no
Área zona (km ²)	Área de la zona en kilómetros cuadrados
Vialidad (km) sin autopistas	Cantidad de kilómetros de vialidad por zona, sin considerar autopistas
Vialidad total (km)	Cantidad de kilómetros de vialidad por zona, considerando autopistas

Variable	Descripción
Total de predios	Total de predios por zona
COMERCIO	Cantidad de predios con destino de uso de suelo final Comercio
DEPORTE Y RECREACION	Cantidad de predios con destino de uso de suelo final Deporte y Recreación
EDUCACION CULTURA	Cantidad de predios con destino de uso de suelo final Educación y cultura
HOTEL MOTEL	Cantidad de predios con destino de uso de suelo final Hotel y Motel
HABITACIONAL	Cantidad de predios con destino de uso de suelo final Habitacional
INDUSTRIA	Cantidad de predios con destino de uso de suelo final Industria
BIENES COMUNES	Cantidad de predios con destino de uso de suelo final Bienes Comunes
BODEGA Y ALMACENAJE	Cantidad de predios con destino de uso de suelo final Bodega y Almacenaje
MINERIA	Cantidad de predios con destino de uso de suelo final Minería
OFICINA	Cantidad de predios con destino de uso de suelo final Oficina
ADMINISTRACION PUBLICA Y DEFENSA	Cantidad de predios con destino de uso de suelo final Administración pública y defensa
CULTO	Cantidad de predios con destino de uso de suelo final culto
SALUD	Cantidad de predios con destino de uso de suelo final Salud
TRANSPORTE Y TELECOMUNICACIONES	Cantidad de predios con destino de uso de suelo final Transporte y Telecomunicaciones
OTROS NO CONSIDERADOS	Cantidad de predios con destino de uso de suelo final Otros no considerados
SITIO ERIAZO	Cantidad de predios con destino de uso de suelo final Sitio Eriazo
ESTACIONAMIENTO	Cantidad de predios con destino de uso de suelo final Estacionamiento
Suma de Avalúo Fiscal de los predios	Suma de avalúo fiscal de los predios
Flujo_vehicular	Flujo de vehículos por zona por día
Veh_km	Suma de vehículo-kilómetro en la zona
Transbordos	Cantidad de transbordos realizados en una zona
Carga_bici	Flujo de bicicletas por zona por día
Carga_cam	Flujo de viajes enteramente a pie por zona por día
Carga_moto	Flujo de motocicletas y/o acompañantes por zona por día
Flujo_buses_ts	Flujo de buses del Transantiago por zona por día
acc_tot_2013	Accidentes totales en el año 2013 por zona
acc_na_2013	Accidentes tipo no atropello en el año 2013 por zona
acc_at_2013	Accidentes tipo atropello en el año 2013 por zona
acc_les_2013	Cantidad de lesionados en todos los accidentes en el año 2013 por zona
acc_fall_2013	Cantidad de fallecidos en todos los accidentes en el año 2013 por zona
In_fv	Logaritmo natural del flujo de vehículos
In_fbici	Logaritmo natural del flujo de bicicletas
In_fp	Logaritmo natural del flujo peatonal
In_fm	Logaritmo natural del flujo de motocicletas
In_fbTS	Logaritmo natural del flujo de buses del Transantiago
In_transbordos	Logaritmo natural de los transbordos

Tabla 6-1: Variables consideradas en la modelación y su descripción.

Fuente: Elaboración propia.

Al buscar qué variables incluir en la modelación, se intenta mejorar la precisión de la estimación con los datos observados. Se intenta generar variables explicativas que resulten estadísticamente significativas en el modelo dado que el resultado será más probablemente estable numéricamente y se puede adecuar mejor a las necesidades de usos futuros, por ejemplo, en la estimación de costos sociales por reducción de accidentes de tránsito.

El proceso de selección de variables implicó un análisis previo de las correlaciones entre ellas, donde se pudo identificar que existían relaciones importantes entre predictores que intuitivamente podrían presentar correspondencias. El resultado se muestra en la Tabla 6-2.

Variables correlacionadas con la primera columna							
Vialidad (km) sin autopistas	Vialidad total (km)	HABITACIONAL					
Vialidad total (km)	Vialidad (km) sin autopistas	HABITACIONAL					
Total de predios	HABITACIONAL	BIENES COMUNES	BODEGA Y ALMACENAJE	ESTACIONAMIENTO	Suma de Avalúo Fiscal de los predios	Flujo_vehicular	veh_km_vehicular
COMERCIO	OFICINA						
DEPORTE Y RECREACION	SITIO ERIAZO	Suma de Avalúo Fiscal de los predios					
EDUCACION CULTURA	HABITACIONAL						
HOTEL MOTEL	ESTACIONAMIENTO						
HABITACIONAL	Vialidad (km) sin autopistas	Vialidad total (km)	Total de predios	EDUCACION CULTURA	BIENES COMUNES	CULTO	
BIENES COMUNES	Total de predios	HABITACIONAL	BODEGA Y ALMACENAJE	ESTACIONAMIENTO	Suma de Avalúo Fiscal de los predios	Flujo_vehicular	
BODEGA Y ALMACENAJE	Total de predios	BIENES COMUNES	ESTACIONAMIENTO	Flujo_vehicular	veh_km_vehicular	Carga_bici	
OFICINA	COMERCIO						
CULTO	HABITACIONAL						
OTROS NO CONSIDERADOS	SITIO ERIAZO						
SITIO ERIAZO	DEPORTE Y RECREACION	OTROS NO CONSIDERADOS	Suma de Avalúo Fiscal de los predios				
ESTACIONAMIENTO	Total de predios	HOTEL MOTEL	BIENES COMUNES	BODEGA Y ALMACENAJE	Suma de Avalúo Fiscal de los predios	Flujo_vehicular	veh_km_vehicular
Suma de Avalúo Fiscal de los predios	Total de predios	DEPORTE Y RECREACION	BIENES COMUNES	SITIO ERIAZO	ESTACIONAMIENTO		
Flujo_vehicular	Total de predios	BIENES COMUNES	BODEGA Y ALMACENAJE	ESTACIONAMIENTO	veh_km_vehicular	Carga_bici	In_fv
veh_km_vehicular	Total de predios	BODEGA Y ALMACENAJE	ESTACIONAMIENTO	Flujo_vehicular	Carga_bici	In_fv	
Carga_bici	BODEGA Y ALMACENAJE	Flujo_vehicular	veh_km_vehicular	Carga_cam	Carga_moto	In_fv	In_fbici
Carga_cam	Carga_bici	In_fp					
Carga_moto	Carga_bici	In_fm					
Flujo_buses_ts	acc_tot_2013	acc_na_2013	acc_at_2013	In_fbTS	In_transbordos		
acc_tot_2013	Flujo_buses_ts	acc_na_2013	acc_at_2013	acc_les_2013			
acc_na_2013	Flujo_buses_ts	acc_tot_2013	acc_at_2013	acc_les_2013			
acc_at_2013	Flujo_buses_ts	acc_tot_2013	acc_na_2013	acc_les_2013			
acc_les_2013	acc_tot_2013	acc_na_2013	acc_at_2013				
In_fv	Flujo_vehicular	veh_km_vehicular	Carga_bici	In_fbTS			
In_fbici	Carga_bici	In_fp					
In_fp	Carga_cam	In_fbici	In_fbTS				
In_fm	Carga_moto						
In_fbTS	Flujo_buses_ts	In_fv	In_fp	In_transbordos			
In_transbordos	Flujo_buses_ts	In_fbTS					

Tabla 6-2: Correlaciones entre las variables.
Fuente: Elaboración propia.

De la Tabla 6-2 las variables que no se incluyen es debido a que no poseen correlación con ninguna otra de las presentadas. Se asume que una correlación es importante cuando el coeficiente de correlación de Pearson es mayor a 0,5, en valor absoluto.

Se puede notar que el uso de suelo habitacional está muy correlacionado con diversas variables: vialidad, cantidad de predios, educación y cultura, bienes comunes y culto. Esto sugiere que se debe simplificar los modelos que incluyan estas variables y considerar en cambio este uso de suelo.

Sobre la correlación del flujo vehicular, también se nota que se correlaciona con algunas variables de uso de suelo, además de la carga de bicicleta por zona. Sin embargo, el logaritmo natural del flujo vehicular no presenta correlaciones con las variables de uso de suelo.

La carga de bicicleta y su logaritmo natural se nota que posee correlación con el flujo vehicular y la carga de caminata, 2 de las variables importantes de los análisis. Es por esta razón que no se considera en los modelos posteriores.

La variable de uso de suelo “oficina” está correlacionada con “comercio”, lo que sugiere intentar realizar comparaciones entre si incluir una o la otra. Esto fue considerado en los modelos y se observó que solo en el caso de accidentes del tipo atropello resulta significativa la variable comercio. La variable oficina si se incluyó a cambio de comercio en los modelos de accidentes totales y los de tipo no atropello por resultar significativa.

Por último, la correlación presente entre el flujo de buses de Transantiago y las variables explicadas sugiere que esta poseerá un alto aporte en la explicación del modelo, tal cual se observará más adelante en los resultados de los modelos.

Los mayores valores de correlación, por sobre 0,8 de coeficiente de Pearson, son entre:

- “Vialidad Total” y “Vialidad sin kilómetros de autopistas”.
- “Total de predios” con “Habitacional” y “Bodega y almacenaje”.
- “Bodega y almacenaje” y “Estacionamiento”.
- “Flujo Vehicular” y “Vehículo kilómetro”.
- “Accidentes totales” con “accidentes no atropello” y “cantidad de lesionados totales”.

Luego del análisis de correlaciones, se procedió a generar los modelos por cada formulación (Normal, Poisson y Binomial Negativo) utilizando la regresión de modelos lineales generalizados en el software SPSS.

Dicho software permite seleccionar las opciones de modelación acorde a lo necesitado. Se presentan en las figuras del anexo 10.7 las opciones utilizadas para la modelación. De esto se destaca que la estimación de parámetros para todas las especificaciones fue realizada mediante el método híbrido, y la estimación del parámetro de escala mediante Chi-cuadrado de Pearson. Además, para la modelación de distribución binomial negativa se considera que el parámetro se debe calcular en vez de ser fijo, al igual que en la modelación Poisson. En todos ellos, como ya se mencionó anteriormente en la formulación matemática de los modelos, consideran una función de enlace logarítmica. Por otro lado, el estadístico para evaluar la significancia de los modelos y los intervalos de confianza es de tipo Wald.

En una primera instancia se incluyeron todas las variables con la intención de ver el efecto agregado. Luego se descartaban variables no significativas y que, al no considerarlas, se generara un mejor ajuste de los modelos. Este ajuste se observa en la devianza y el indicador AIC.

Tras obtener un grupo de variables significativas en la modelación, se consideraban las correlaciones y, si era necesario, se descartaban las que tuvieran un alto coeficiente de Pearson. Así se probaron varias formulaciones para terminar concluyendo que las variables que son más significativas y que se repiten en todos los tipos de modelos, sin considerar las variables dummy de pertenencia a una comuna, son: Habitacional, Oficina, Flujo vehicular, Flujo de buses y Flujo peatonal.

Cabe destacar que el análisis del factor de inflación de la varianza (VIF por sus siglas en inglés) es necesario al momento de observar multicolinealidad. En los modelos presentados ninguno se vio afectado por este fenómeno, es decir, en ningún caso existió un valor del estadístico VIF mayor a 3. Es más, en los modelos considerados, el máximo valor de este estadístico fue de 2,55 y el mínimo fue de 1,01 como se muestra en la Tabla 6-3. El cálculo del factor de inflación de la varianza se realiza para las variables incluidas en cada modelación, según tipo de variable explicada.

Se consideraron 2 tipos de resultados de la modelación. Uno de ellos consta de variables que en todos los modelos resultaron significativas luego del proceso de selección (Modelo 1), y otra modelación, que incluye más variables, donde en al menos una modelación resultaron significativas (Modelo 2).

Variables	VIF por modelo				
	Accidentes totales	Accidentes no atropello	Accidentes atropello	Lesionados	Fallecidos
DummyStgo	1,291	1,455			
DummyCerrillos	1,056	1,094	1,041	1,027	
DummyCerroNavia	1,055	1,075			
DummyConchali	1,079	1,110			
DummyElBosque	1,091	1,106		1,058	
DummyEstacionCentral	1,083	1,130			
DummyIndependencia	1,060	1,087			
DummyLaCisterna	1,086	1,108			
DummyLaFlorida	1,159	1,248	1,087		
DummyLaGranja		1,088		1,032	
DummyLaPintana				1,055	1,037
DummyLaReina		1,170		1,048	
DummyLasCondes	1,221	1,410	1,224		
DummyLoEspejo	1,047	1,062			
DummyLoPrado	1,085	1,093	1,047		
DummyMacul	1,074	1,136	1,033		
DummyMaipu	1,138	1,215	1,105	1,060	
DummyÑuñoa		1,244		1,065	
DummyPAC				1,030	
DummyPeñalolen	1,094	1,150	1,049		
DummyProvidencia	1,324	1,456			
DummyPudahuel	1,084	1,124			
DummyQtaNormal				1,037	1,021
DummyRecoleta	1,102	1,147	1,046	1,040	
DummyRenca	1,058	1,086			1,011
DummySanJoaquin	1,081	1,119	1,035		1,014
DummySanMiguel	1,071	1,094	1,045	1,028	
DummySanRamon	1,059			1,030	
DummyVitacura		1,276		1,119	
DummyPuenteAlto				1,072	1,036
Vialidad (km) sin autopistas			1,399		
HABITACIONAL	1,269	1,287		1,174	1,076
OFICINA	1,217	1,222			
COMERCIO			1,279		
INDUSTRIA			1,116		
CULTO			1,239		
SALUD	1,160	1,164	1,113		
Suma de Avalúo Fiscal de los predios			1,511		
In_fv	2,117	2,548	2,096	1,944	1,539
In_fp	1,503	1,574	1,525	1,501	
In_fm			1,420		
In_fbTS	2,077	2,109	1,983	1,975	1,440
In_transbordos	1,653	1,672	1,656	1,581	

Tabla 6-3: Factor de inflación de la varianza (VIF) por cada modelación.
Fuente: Elaboración propia en base a modelación.

6.3 Modelos lineales generalizados Normal con función de enlace logarítmica

Sobre la forma funcional de los modelos a considerar, la revisión bibliográfica sugiere una formulación como la presentada en el apartado 3 mediante la Ecuación 13.

Esto indica que se puede considerar una regresión con el predictor lineal, el cual, en este caso de los modelos de regresión, se asume que sigue una distribución normal. Se llaman log lineal por la función logarítmica a un costado de la igualdad, y la linealidad del otro.

Cabe destacar que mediante la aplicación de una transformación adecuada se podrían normalizar los datos, particularmente la variable explicada. Esto se justifica debido a la propuesta establecida por George E. P. Box y David Cox (Box & Cox, 1981) en donde plantean una familia de transformaciones potenciales usadas para corregir sesgos en la distribución de errores, corregir varianzas desiguales y para corregir la no linealidad en las relaciones.

El resultado de ese proceso de transformaciones indica que la más adecuada para las variables explicativas de este modelo es la logarítmica. Entonces, debido a que la función de enlace es ella misma, los datos serán naturalmente normalizados mediante la formulación matemática de los modelos y, por lo tanto, se justifica la utilización de un modelo de regresión lineal. Esto resulta en una forma simple de modelación en el que se considera una variable explicada parecida a una distribución normal, junto con diversos factores explicativos considerados bajo los supuestos de modelamiento.

Sobre la modelación de regresión lineal, los supuestos detrás son los siguientes:

- Independencia: Se asume que los residuos serán independientes.
- Homocedasticidad: La homocedasticidad o igualdad de varianzas de los residuos implica que la variación de estos sea uniforme en todo el rango de valores de los pronósticos.
- Normalidad de los residuos: La normalidad de los residuos tipificados se puede contrastar mediante la prueba de Kolmogorov-Smirnov o mediante análisis gráfico del histograma de la variable dependiente, junto con el gráfico de probabilidad normal P-P que representa las proporciones acumuladas de la variable esperada respecto a las proporciones acumuladas de la variable observada.
- No-colinealidad: La inexistencia de colinealidad entre las variables independientes es parte importante de los supuestos de las regresiones lineales. Esta se analiza mediante el análisis de correlaciones del estadístico de Pearson, tal cual se presentó anteriormente.
- Linealidad: Si no se tiene linealidad se dice que se tiene un error de especificación.

Los resultados de los modelos considerando una distribución normal son presentados en la Tabla 6-4, Tabla 6-5, Tabla 6-6, Tabla 6-7 y Tabla 6-8.

Parámetros	Acc totales 1		Acc totales 2	
	B	Sig	B	Sig
<i>(Intersección)</i>	-2,96 E+00	0,000	-2,81 E+00	0,000
<i>DummyStgo</i>	-3,57 E-01	0,000	-3,69 E-01	0,000
<i>DummyCerrillos</i>	-1,16 E+00	0,012	-1,21 E+00	0,006
<i>DummyCerroNavia</i>			-2,15 E-01	0,219
<i>DummyConchali</i>	-6,14 E-01	0,011	-6,92 E-01	0,002
<i>DummyElBosque</i>			-3,90 E-01	0,053
<i>DummyEstacionCentral</i>	-8,59 E-01	0,000	-9,30 E-01	0,000
<i>DummyIndependencia</i>	-5,58 E-01	0,015	-6,41 E-01	0,003
<i>DummyLaCisterna</i>			-2,58 E-01	0,042
<i>DummyLaFlorida</i>			-6,50 E-02	0,336
<i>DummyLasCondes</i>	-8,49 E-01	0,000	-8,18 E-01	0,000
<i>DummyLoEspejo</i>			-4,14 E-01	0,117
<i>DummyLoPrado</i>			-5,36 E-01	0,007
<i>DummyMacul</i>	-5,55 E-01	0,001	-6,12 E-01	0,000
<i>DummyMaipu</i>	-1,79 E-01	0,024	-3,09 E-01	0,000
<i>DummyPeñalolen</i>	-5,70 E-01	0,001	-6,64 E-01	0,000
<i>DummyProvidencia</i>	-7,53 E-01	0,000	-9,06 E-01	0,000
<i>DummyPudahuel</i>			-3,64 E-01	0,025
<i>DummyRecoleta</i>	-7,09 E-01	0,000	-7,75 E-01	0,000
<i>DummyRenca</i>			-2,59 E-01	0,118
<i>DummySanJoaquin</i>	-3,86 E-01	0,039	-4,49 E-01	0,011
<i>DummySanMiguel</i>	-1,23 E+00	0,000	-1,35 E+00	0,000
<i>DummySanRamon</i>			2,47 E-01	0,038
<i>HABITACIONAL</i>	1,57 E-05	0,004	1,01 E-05	0,041
<i>OFICINA</i>			-8,80 E-05	0,033
<i>SALUD</i>			8,95 E-03	0,000
<i>ln_fv</i>	3,25 E-01	0,000	2,90 E-01	0,000
<i>ln_fp</i>	7,71 E-02	0,002	8,53 E-02	0,000
<i>ln_fbTS</i>	3,24 E-01	0,000	3,52 E-01	0,000
<i>ln_transbordos</i>	4,63 E-02	0,000	4,54 E-02	0,000
<i>Desviianza</i>	163.606,2		142.122,4	
<i>Logaritmo de verosimilitud</i>	-82.479,4		-71.737,5	
<i>Criterio de información Akaike (AIC)</i>	164.996,9		143.535,1	

Tabla 6-4: Resultados de la modelación Normal para accidentes totales.
Fuente: Elaboración propia.

En los resultados presentados en la Tabla 6-4 la variable explicada corresponde a la cantidad de accidentes totales por zona. Las variables explicativas que se consideran entre los modelos tipo 1 y 2 son debido a la consideración de las variables en los modelos donde todas son significativas a un 90 % de confianza (modelos 1) y la modelación donde en al menos una formulación (Normal, Poisson o Binomial Negativa) las variables son significativas (modelos 2). Esto rige para todos los modelos presentados en este trabajo.

Los estimadores encontrados para las variables dummies y la intersección son todos negativos, excepto para el modelo tipo 2 en la variable dummy de San Ramón. En cuanto a las variables de uso de suelo, flujos y transbordos resultan todos los estimadores positivos, excepto el de oficina en el modelo tipo 2.

Parámetros	Acc no atropello 1		Acc no atropello 2	
	B	Sig	B	Sig
<i>(Intersección)</i>	-3,27 E+00	0,000	-3,35 E+00	0,000
<i>DummyStgo</i>	-3,35 E-01	0,000	-5,84 E-01	0,000
<i>DummyCerrillos</i>	-1,14 E+00	0,014	-1,35 E+00	0,002
<i>DummyCerroNavia</i>			-3,12 E-01	0,082
<i>DummyConchali</i>	-6,12 E-01	0,014	-8,23 E-01	0,000
<i>DummyElBosque</i>			-5,44 E-01	0,016
<i>DummyEstacionCentral</i>	-9,98 E-01	0,000	-1,23 E+00	0,000
<i>DummyIndependencia</i>	-5,85 E-01	0,015	-7,95 E-01	0,000
<i>DummyLaCisterna</i>			-3,27 E-01	0,009
<i>DummyLaFlorida</i>			-1,57 E-01	0,022
<i>DummyLaGranja</i>			-3,38 E-01	0,012
<i>DummyLaReina</i>			-2,85 E-01	0,013
<i>DummyLasCondes</i>	-8,16 E-01	0,000	-1,06 E+00	0,000
<i>DummyLoEspejo</i>			-4,60 E-01	0,070
<i>DummyLoPrado</i>			-6,61 E-01	0,002
<i>DummyMacul</i>	-5,32 E-01	0,001	-7,72 E-01	0,000
<i>DummyMaipu</i>			-3,77 E-01	0,000
<i>DummyÑuñoa</i>			-3,38 E-01	0,000
<i>DummyPeñalolen</i>	-5,61 E-01	0,002	-8,08 E-01	0,000
<i>DummyProvidencia</i>	-7,62 E-01	0,000	-1,18 E+00	0,000
<i>DummyPudahuel</i>			-4,19 E-01	0,008
<i>DummyRecoleta</i>	-7,51 E-01	0,000	-9,37 E-01	0,000
<i>DummyRenca</i>			-3,82 E-01	0,023
<i>DummySanJoaquin</i>	-3,32 E-01	0,067	-5,34 E-01	0,002
<i>DummySanMiguel</i>	-1,15 E+00	0,000	-1,42 E+00	0,000
<i>DummyVitacura</i>			-3,45 E-01	0,002
<i>HABITACIONAL</i>	1,48 E-05	0,008	1,17 E-05	0,019
<i>OFICINA</i>	-9,78 E-05	0,039	-1,16 E-04	0,019
<i>SALUD</i>			7,98 E-03	0,000
<i>ln_fv</i>	3,53 E-01	0,000	3,84 E-01	0,000
<i>ln_fp</i>	7,87 E-02	0,002	6,39 E-02	0,008
<i>ln_fbTS</i>	3,13 E-01	0,000	3,34 E-01	0,000
<i>ln_transbordos</i>	4,28 E-02	0,000	3,73 E-02	0,000
<i>Desvianza</i>	132.503,7		112.228,2	
<i>Logaritmo de verosimilitud</i>	-66.928,2		-56.790,5	
<i>Criterio de información Akaike (AIC)</i>	133.894,4		113.646,9	

Tabla 6-5: Resultados de la modelación Normal para accidentes tipo no atropello.
Fuente: Elaboración propia.

Los estimadores en la Tabla 6-5 para las variables dummy y la intersección son negativos, al igual que el modelo de accidentes totales. Las variables de flujo, transbordos y uso de suelo tienen estimadores positivos, excepto la variable oficina.

Todas las variables incluidas son significativas por sobre el 90% de significancia. El mayor valor es 0,082 para la variable dummy de Cerro Navia.

Parámetros	Acc atropello 1		Acc atropello 2	
	B	Sig	B	Sig
(Intersección)	-4,64 E+00	0,000	-5,31 E+00	0,000
DummyCerrillos			-1,11 E+00	0,109
DummyLaFlorida	-5,59 E-01	0,001	-5,86 E-01	0,000
DummyLasCondes	-6,41 E-01	0,000	-3,12 E-01	0,060
DummyLoPrado			-4,37 E-01	0,036
DummyMacul	-6,88 E-01	0,031	-6,89 E-01	0,026
DummyMaipu	-8,78 E-01	0,000	-9,36 E-01	0,000
DummyPeñalolen	-4,73 E-01	0,041	-3,92 E-01	0,090
DummyRecoleta			-6,04 E-01	0,004
DummySanJoaquin	-7,97 E-01	0,094	-8,06 E-01	0,076
DummySanMiguel	-1,80 E+00	0,096	-1,84 E+00	0,081
Vialidadkmsinautopistas	7,55 E-03	0,000	8,46 E-03	0,000
COMERCIO	1,21 E-04	0,001	3,78 E-04	0,000
INDUSTRIA	-3,45 E-03	0,001	-2,89 E-03	0,003
CULTO	7,35 E-03	0,038	1,10 E-02	0,001
SALUD			7,18 E-03	0,000
SumadeAvalúoFiscaldelospredios			-8,43 E-13	0,000
In_fv	1,82 E-01	0,000	2,54 E-01	0,000
In_fp	1,19 E-01	0,000	1,96 E-01	0,000
In_fm			-6,56 E-02	0,000
In_fbTS	3,55 E-01	0,000	3,10 E-01	0,000
In_transbordos	7,88 E-02	0,000	7,99 E-02	0,000
Desvianza	4.763,7		4.252,6	
Logaritmo de verosimilitud	-3.058,2		-2.802,6	
Criterio de información Akaike (AIC)	6.148,4		5.649,3	

Tabla 6-6: Resultados de la modelación Normal para accidentes tipo atropello.

Fuente: Elaboración propia.

Todos los estimadores en la Tabla 6-6 son negativos para las variables dummies de comunas y la intersección. Por el contrario, las variables de vialidad sin autopistas, comercio, culto, salud, flujo de automóvil (In_fv), flujo peatonal (In_fp), flujo de buses del Transantiago (In_fbTS) y los transbordos, poseen estimadores que son positivos. Los coeficientes de industria, avalúo fiscal de predios y el flujo de motocicletas son negativos.

Todas las variables resultan significativas por sobre un 90%, a excepción de la dummy de Cerrillos, en que su significancia es de 0,109.

Parámetros	Lesionados 1		Lesionados 2	
	B	Sig	B	Sig
(Intersección)	-3,28 E+00	0,000	-2,88 E+00	0,000
DummyCerrillos			-1,50 E+00	0,172
DummyElBosque	6,89 E-01	0,000	6,06 E-01	0,000
DummyLaGranja	9,83 E-01	0,000	9,35 E-01	0,000
DummyLaPintana	9,14 E-01	0,000	8,22 E-01	0,000
DummyLaReina	4,16 E-01	0,002	3,68 E-01	0,005
DummyMaipu			-2,06 E-01	0,147
DummyÑuñoa	7,39 E-01	0,000	6,91 E-01	0,000
DummyPAC	6,57 E-01	0,000	6,10 E-01	0,000
DummyQtaNormal	1,03 E+00	0,000	9,71 E-01	0,000
DummyRecoleta			-4,32 E-01	0,094
DummySanMiguel	-1,40 E+00	0,058	-1,44 E+00	0,049
DummySanRamon			-4,66 E-01	0,292
DummyVitacura	4,66 E-01	0,000	4,48 E-01	0,000
DummyPuenteAlto	6,65 E-01	0,000	5,90 E-01	0,000
HABITACIONAL			7,02 E-06	0,232
ln_fv	3,04 E-01	0,000	2,75 E-01	0,000
ln_fp	5,95 E-02	0,017	5,23 E-02	0,038
ln_fbTS	2,97 E-01	0,000	2,91 E-01	0,000
ln_transbordos	4,08 E-02	0,001	4,43 E-02	0,000
Desviación	79.414,8		77.440,4	
Logaritmo de verosimilitud	-40.383,7		-39.396,5	
Criterio de información Akaike (AIC)	80.797,4		78.833,1	

Tabla 6-7: Resultados de la modelación Normal para lesionados.
Fuente: Elaboración propia.

Los estimadores positivos de las variables dummy son todos los incluidos para los modelos de la Tabla 6-7, excepto Cerrillos, Maipú, Recoleta, San Miguel y San Ramón. El resto de las variables posee estimadores positivos.

Las significancias de las variables dummy de Cerrillos, Maipú, San Ramón y la variable de uso de suelo Habitacional son mayores a 0,1, lo que indica que no son significativas a un 90% de confianza, es decir, su estimador no es confiablemente distinto de cero.

Parámetros	Fallecidos 1		Fallecidos 2	
	B	Sig	B	Sig
(Intersección)	-5,73 E+00	0,000	-6,16 E+00	0,000
DummyLaPintana	1,13 E+00	0,000	1,36 E+00	0,000
DummyQtaNormal			6,34 E-01	0,051
DummyRenca	8,83 E-01	0,001	1,02 E+00	0,000
DummySanJoaquin	8,84 E-01	0,000	9,47 E-01	0,000
DummyPuenteAlto	6,99 E-01	0,000	8,61 E-01	0,000
HABITACIONAL	3,84 E-05	0,004	3,49 E-05	0,011
ln_fv			1,39 E-01	0,078
ln_fbTS	6,19 E-01	0,000	4,85 E-01	0,000
Desviación	291,0		289,0	
Logaritmo de verosimilitud	-821,9		-820,8	
Criterio de información Akaike (AIC)	1.657,7		1.659,7	

Tabla 6-8: Resultados de la modelación Normal para fallecidos.
Fuente: Elaboración propia.

Todas las variables son significativas y positivas (excepto la intersección) en la Tabla 6-8.

6.4 Modelos lineales generalizados Poisson con función de enlace logarítmica

El principal supuesto detrás de la regresión Poisson es que la media de la frecuencia de accidentes debe ser igual a su varianza. Para analizar lo anterior se muestran los resultados en la Tabla 6-9 para distintos tipos de accidentes y frecuencias consideradas en 5 años.

		2011	2012	2013	2014	2015
Accidentes totales	Media	23,91	23,66	24,90	25,32	26,24
	Varianza	471,15	423,88	475,86	596,17	611,12
Accidentes tipo no atropello	Media	20,93	20,75	22,04	22,38	23,29
	Varianza	363,66	323,75	376,47	478,97	494,18
Accidentes tipo atropello	Media	3,85	3,88	3,75	3,83	3,75
	Varianza	13,68	11,56	11,81	11,57	10,60
Cantidad de lesionados	Media	14,92	14,66	15,14	14,76	14,99
	Varianza	203,92	191,05	207,52	211,50	204,66
Cantidad de fallecidos	Media	0,30	0,29	0,34	0,34	0,32
	Varianza	0,36	0,41	0,44	0,45	0,44

Tabla 6-9: Resumen de Media y Varianza para distintos tipos de accidentes y frecuencias por zona.
Fuente: Elaboración propia en base a SIEC-2

De lo anterior se puede notar que la varianza difiere bastante de la media en todos los años y en todos los casos, incluso para la frecuencia de fallecidos por zona que es donde se parecen más dichos valores. Esto indica que existe sobredispersión en los datos (Varianza mayor que la media).

De igual forma se estiman modelos lineales generalizados bajo el supuesto de que el error distribuye Poisson y la función de enlace es logarítmica dada la especificación planteada en la Ecuación 11 en el apartado 3, considerando los accidentes del año 2013 por zona, según el total de accidentes, accidentes tipo atropello, accidentes tipo no atropello, cantidad de lesionados y cantidad de fallecidos.

El procedimiento para estimar estos modelos es similar al realizado para el caso de regresión lineal multivariada y se presentan los mismos tipos de modelos 1 y 2.

Los resultados de la modelación son presentados en la Tabla 6-10, Tabla 6-11, Tabla 6-12, Tabla 6-13 y Tabla 6-14.

Parámetros	Acc totales 1		Acc totales 2	
	B	Sig	B	Sig
(Intersección)	-3,20 E+00	0,000	-3,09 E+00	0,000
DummyStgo	-2,82 E-01	0,000	-3,13 E-01	0,000
DummyCerrillos	-1,26 E+00	0,000	-1,32 E+00	0,000
DummyCerroNavia			-2,46 E-01	0,128
DummyConchali	-5,89 E-01	0,001	-6,78 E-01	0,000
DummyElBosque			-4,34 E-01	0,011
DummyEstacionCentral	-8,87 E-01	0,000	-9,70 E-01	0,000
DummyIndependencia	-6,16 E-01	0,001	-7,12 E-01	0,000
DummyLaCisterna			-2,12 E-01	0,103
DummyLaFlorida			-1,73 E-01	0,036
DummyLasCondes	-8,08 E-01	0,000	-8,11 E-01	0,000
DummyLoEspejo			-4,12 E-01	0,046
DummyLoPrado			-4,07 E-01	0,019
DummyMacul	-5,26 E-01	0,000	-5,96 E-01	0,000
DummyMaipu	-1,83 E-01	0,033	-2,90 E-01	0,001
DummyPeñalolen	-5,32 E-01	0,000	-6,21 E-01	0,000
DummyProvidencia	-7,25 E-01	0,000	-7,82 E-01	0,000
DummyPudahuel			-3,10 E-01	0,020
DummyRecoleta	-6,81 E-01	0,000	-7,64 E-01	0,000
DummyRenca			-2,78 E-01	0,067
DummySanJoaquin	-3,70 E-01	0,015	-4,49 E-01	0,002
DummySanMiguel	-1,12 E+00	0,000	-1,24 E+00	0,000
DummySanRamon			2,46 E-01	0,076
HABITACIONAL	1,78 E-05	0,005	1,44 E-05	0,018
OFICINA			-9,97 E-05	0,020
SALUD			6,52 E-03	0,000
ln_fv	3,42 E-01	0,000	3,12 E-01	0,000
ln_fp	7,61 E-02	0,000	8,70 E-02	0,000
ln_fbTS	3,39 E-01	0,000	3,61 E-01	0,000
ln_transbordos	3,97 E-02	0,000	4,00 E-02	0,000
Desvianza	4979,9		4623,8	
Chi-cuadrado de Pearson	5364,8		4842,9	
Logaritmo de verosimilitud	-4180,3		-4002,3	
Criterio de información Akaike (AIC)	8398,7		8064,6	

Tabla 6-10: Resultados de la modelación Poisson para accidentes totales.
Fuente: Elaboración propia.

En la Tabla 6-10 se puede notar que las variables dummy son todas negativas, excepto en la comuna de San Ramón. También es positivo el estimador de la variable oficina. El resto de las variables poseen coeficientes positivos.

Todas las variables son significativas, excepto la dummy de Cerro Navia y de La Cisterna.

Parámetros	Acc no atropello 1		Acc no atropello 2	
	B	Sig	B	Sig
(Intersección)	-3,48 E+00	0,000	-3,39 E+00	0,000
DummyStgo	-2,63 E-01	0,000	-4,59 E-01	0,000
DummyCerrillos	-1,23 E+00	0,000	-1,41 E+00	0,000
DummyCerroNavia			-3,27 E-01	0,048
DummyConchali	-6,01 E-01	0,001	-7,91 E-01	0,000
DummyElBosque			-5,68 E-01	0,002
DummyEstacionCentral	-1,03 E+00	0,000	-1,22 E+00	0,000
DummyIndependencia	-6,39 E-01	0,001	-8,35 E-01	0,000
DummyLaCisterna			-2,63 E-01	0,045
DummyLaFlorida			-2,37 E-01	0,006
DummyLaGranja			-2,11 E-01	0,128
DummyLaReina			-1,94 E-01	0,106
DummyLasCondes	-7,78 E-01	0,000	-9,68 E-01	0,000
DummyLoEspejo			-4,52 E-01	0,028
DummyLoPrado			-5,06 E-01	0,005
DummyMacul	-5,00 E-01	0,000	-6,98 E-01	0,000
DummyMaipu			-3,34 E-01	0,000
DummyÑuñoa			-2,73 E-01	0,005
DummyPeñalolen	-5,31 E-01	0,000	-7,28 E-01	0,000
DummyProvidencia	-7,50 E-01	0,000	-1,02 E+00	0,000
DummyPudahuel			-3,70 E-01	0,006
DummyRecoleta	-7,33 E-01	0,000	-9,16 E-01	0,000
DummyRenca			-3,70 E-01	0,017
DummySanJoaquin	-3,22 E-01	0,033	-5,06 E-01	0,001
DummySanMiguel	-1,03 E+00	0,000	-1,26 E+00	0,000
DummyVitacura			-2,15 E-01	0,087
HABITACIONAL	1,60 E-05	0,012	1,43 E-05	0,021
OFICINA	-1,05 E-04	0,026	-1,25 E-04	0,007
SALUD			6,04 E-03	0,000
ln_fv	3,71 E-01	0,000	3,72 E-01	0,000
ln_fp	7,32 E-02	0,001	7,48 E-02	0,000
ln_fbTS	3,28 E-01	0,000	3,39 E-01	0,000
ln_transbordos	3,66 E-02	0,000	3,44 E-02	0,000
Desviianza	4556,5		4199,1	
Chi-cuadrado de Pearson	4873,4		4360,0	
Logaritmo de verosimilitud	-3918,2		-3739,5	
Criterio de información Akaike (AIC)	7874,4		7545,0	

Tabla 6-11: Resultados de la modelación Poisson para accidentes tipo no atropello.

Fuente: Elaboración propia.

En la Tabla 6-11 todos los estimadores de las variables dummy son negativos, además de la variable oficina. El resto de las variables presentan estimadores positivos.

Las variables dummy de La Granja y La Reina no son significativas a un 90% de confianza. El resto sí lo son.

Parámetros	Acc atropello 1		Acc atropello 2	
	B	Sig	B	Sig
(Intersección)	-5,06 E+00	0,000	-5,35 E+00	0,000
DummyCerrillos			-1,29 E+00	0,006
DummyLaFlorida	-4,95 E-01	0,001	-5,35 E-01	0,000
DummyLasCondes	-6,12 E-01	0,000	-4,68 E-01	0,005
DummyLoPrado			-1,79 E-01	0,426
DummyMacul	-5,44 E-01	0,025	-5,24 E-01	0,027
DummyMaipu	-7,45 E-01	0,000	-7,87 E-01	0,000
DummyPeñalolen	-4,35 E-01	0,030	-3,83 E-01	0,052
DummyRecoleta			-3,42 E-01	0,071
DummySanJoaquin	-6,25 E-01	0,026	-6,22 E-01	0,024
DummySanMiguel	-1,73 E+00	0,000	-1,74 E+00	0,000
Vialidadkmsinautopistas	8,19 E-03	0,000	8,91 E-03	0,000
COMERCIO	1,98 E-04	0,006	2,60 E-04	0,002
INDUSTRIA	-3,02 E-03	0,000	-2,33 E-03	0,005
CULTO	1,14 E-02	0,005	9,98 E-03	0,014
SALUD			5,02 E-03	0,045
SumadeAvalúoFiscaldelospredios			-2,21 E-13	0,083
ln_fv	1,83 E-01	0,000	2,30 E-01	0,000
ln_fp	1,26 E-01	0,000	1,58 E-01	0,000
ln_fm			-5,18 E-02	0,000
ln_fbTS	4,04 E-01	0,000	3,91 E-01	0,000
ln_transbordos	6,56 E-02	0,000	6,02 E-02	0,000
Desvianza	1413,8		1350,9	
Chi-cuadrado de Pearson	1447,4		1362,9	
Logaritmo de verosimilitud	-1533,2		-1501,8	
Criterio de información Akaike (AIC)	3098,4		3047,6	

Tabla 6-12: Resultados de la modelación Poisson para accidentes tipo atropello.
Fuente: Elaboración propia.

La Tabla 6-12 muestra que los estimadores de las variables dummy de comunas, junto con la variable industria, suma del avalúo fiscal de los predios y el flujo de motocicletas son negativas. Para el resto, son positivas.

Todas las variables resultan significativas, a excepción de la variable dummy Lo Prado.

Parámetros	Lesionados 1		Lesionados 2	
	B	Sig	B	Sig
(Intersección)	-3,74 E+00	0,000	-3,37 E+00	0,000
DummyCerrillos			-1,53 E+00	0,000
DummyElBosque	6,97 E-01	0,000	6,07 E-01	0,000
DummyLaGranja	1,06 E+00	0,000	1,00 E+00	0,000
DummyLaPintana	8,59 E-01	0,000	7,66 E-01	0,000
DummyLaReina	4,33 E-01	0,002	3,65 E-01	0,007
DummyMaipu			-2,10 E-01	0,083
DummyÑuñoa	7,29 E-01	0,000	6,59 E-01	0,000
DummyPAC	6,43 E-01	0,000	5,97 E-01	0,000
DummyQtaNormal	1,03 E+00	0,000	9,78 E-01	0,000
DummyRecoleta			-4,20 E-01	0,024
DummySanMiguel	-1,43 E+00	0,000	-1,47 E+00	0,000
DummySanRamon			-5,29 E-01	0,074
DummyVitacura	5,12 E-01	0,000	4,84 E-01	0,000
DummyPuenteAlto	5,94 E-01	0,000	4,97 E-01	0,000
HABITACIONAL			1,35 E-05	0,060
In_fv	3,22 E-01	0,000	2,95 E-01	0,000
In_fp	6,98 E-02	0,003	5,93 E-02	0,012
In_fbTS	3,29 E-01	0,000	3,29 E-01	0,000
In_transbordos	3,43 E-02	0,003	3,54 E-02	0,002
Desvianza	4274,0		4052,0	
Chi-cuadrado de Pearson	4575,9		4340,6	
Logaritmo de verosimilitud	-3599,3		-3488,3	
Criterio de información Akaike (AIC)	7228,6		7016,5	

Tabla 6-13: Resultados de la modelación Poisson para lesionados.

Fuente: Elaboración propia.

En la Tabla 6-13 se nota que las variables dummy con estimadores negativos son las de Cerrillos, Maipú, Recoleta, San Miguel y San Ramón. Todo el resto de las variables son positivas.

Sobre la significancia, se observa que todas las variables presentadas en la Tabla 6-13 son significativas.

Parámetros	Fallecidos 1		Fallecidos 2	
	B	Sig	B	Sig
<i>(Intersección)</i>	-5,78 E+00	0,000	-6,22 E+00	0,000
<i>DummyLaPintana</i>	1,02 E+00	0,001	1,19 E+00	0,000
<i>DummyQtaNormal</i>			5,75 E-01	0,134
<i>DummyRenca</i>	7,56 E-01	0,047	8,29 E-01	0,030
<i>DummySanJoaquin</i>	7,68 E-01	0,027	8,28 E-01	0,017
<i>DummyPuenteAlto</i>	6,72 E-01	0,007	7,61 E-01	0,003
<i>HABITACIONAL</i>	4,05 E-05	0,037	3,83 E-05	0,053
<i>ln_fv</i>			1,11 E-01	0,183
<i>ln_fbTS</i>	6,28 E-01	0,000	5,37 E-01	0,000
<i>Desviianza</i>	660,4		656,3	
<i>Chi-cuadrado de Pearson</i>	898,3		888,7	
<i>Logaritmo de verosimilitud</i>	-523,8		-521,8	
<i>Criterio de información Akaike (AIC)</i>	1061,7		1061,5	

Tabla 6-14: Resultados de la modelación Poisson para fallecidos.

Fuente: Elaboración propia.

En los resultados de la modelación de Poisson para la cantidad de fallecidos por zona de la Tabla 6-14, se observa que todas las variables poseen estimadores positivos. Además, las únicas 2 variables que no son significativas son la correspondiente a la dummy de pertenencia de zona en la comuna de Quinta Normal y el flujo de automóvil.

6.5 Modelos lineales generalizados Binomial Negativa con función de enlace logarítmica

La regresión mediante este tipo de modelos considera casi siempre datos de conteo que poseen sobredispersión. Describe la probabilidad de ocurrencia de un número entero de veces que ocurre un fenómeno en donde la varianza depende de la media y del parámetro de sobredispersión, como se mostró en la Ecuación 12.

Los resultados se notan en la Tabla 6-15, Tabla 6-16, Tabla 6-17, Tabla 6-18 y Tabla 6-19.

Parámetros	Acc totales 1		Acc totales 2	
	B	Sig.	B	Sig.
(Intersección)	-3,17 E+00	0,000	-3,10 E+00	0,000
DummyStgo	-1,84 E-01	0,047	-2,36 E-01	0,013
DummyCerrillos	-1,30 E+00	0,000	-1,37 E+00	0,000
DummyCerroNavia			-2,78 E-01	0,091
DummyConchali	-5,55 E-01	0,000	-6,44 E-01	0,000
DummyElBosque			-4,60 E-01	0,003
DummyEstacionCentral	-8,83 E-01	0,000	-9,67 E-01	0,000
DummyIndependencia	-6,60 E-01	0,000	-7,53 E-01	0,000
DummyLaCisterna			-1,38 E-01	0,350
DummyLaFlorida			-2,45 E-01	0,012
DummyLasCondes	-7,08 E-01	0,000	-7,45 E-01	0,000
DummyLoEspejo			-3,88 E-01	0,040
DummyLoPrado			-3,21 E-01	0,076
DummyMacul	-4,92 E-01	0,001	-5,66 E-01	0,000
DummyMaipu	-1,75 E-01	0,080	-2,74 E-01	0,006
DummyPeñalolen	-4,92 E-01	0,000	-5,79 E-01	0,000
DummyProvidencia	-7,03 E-01	0,000	-7,32 E-01	0,000
DummyPudahuel			-2,48 E-01	0,056
DummyRecoleta	-6,45 E-01	0,000	-7,27 E-01	0,000
DummyRenca			-2,29 E-01	0,133
DummySanJoaquin	-3,50 E-01	0,015	-4,30 E-01	0,003
DummySanMiguel	-1,02 E+00	0,000	-1,13 E+00	0,000
DummySanRamon			2,63 E-01	0,112
HABITACIONAL	1,98 E-05	0,014	1,89 E-05	0,018
OFICINA			-9,90 E-05	0,077
SALUD			3,56 E-03	0,166
In_fv	3,56 E-01	0,000	3,36 E-01	0,000
In_fp	7,79 E-02	0,000	8,48 E-02	0,000
In_fbTS	3,12 E-01	0,000	3,32 E-01	0,000
In_transbordos	3,61 E-02	0,000	3,77 E-02	0,000
Sobredispersión	2,63 E-01		2,46 E-01	
Desviianza	811,8		815,7	
Chi-cuadrado de Pearson	830,5		820,0	
Logaritmo de verosimilitud	-2748,5		-2730,3	
Criterio de información Akaike (AIC)	5536,9		5522,7	

Tabla 6-15: Resultados de la modelación Binomial Negativa para accidentes totales.
Fuente: Elaboración propia.

Los resultados de la Tabla 6-15 muestran coeficientes de variables dummy negativas para todos los casos, a excepción de uno: el de la comuna de San Ramón. Además, los estimadores de la variable oficina también es negativo.

Las variables que no son significativas a un 90% de confianza son la dummy de La Cisterna, dummy de Renca, dummy de San Ramón y el uso de suelo Salud.

Se destaca que también se presenta el valor de la sobredispersión que se estima por el modelo mediante la herramienta SPSS. Difieren entre ambos en poco más de 2 centésimas.

Parámetros	Acc no atropello 1		Acc no atropello 2	
	B	Sig.	B	Sig.
<i>(Intersección)</i>	-3,44 E+00	0,000	-3,36 E+00	0,000
<i>DummyStgo</i>	-1,76 E-01	0,059	-3,24 E-01	0,001
<i>DummyCerrillos</i>	-1,28 E+00	0,000	-1,41 E+00	0,000
<i>DummyCerroNavia</i>			-3,26 E-01	0,051
<i>DummyConchali</i>	-5,84 E-01	0,000	-7,35 E-01	0,000
<i>DummyElBosque</i>			-5,64 E-01	0,000
<i>DummyEstacionCentral</i>	-1,03 E+00	0,000	-1,18 E+00	0,000
<i>DummyIndependencia</i>	-6,84 E-01	0,000	-8,40 E-01	0,000
<i>DummyLaCisterna</i>			-1,66 E-01	0,267
<i>DummyLaFlorida</i>			-2,76 E-01	0,007
<i>DummyLaGranja</i>			-8,36 E-03	0,958
<i>DummyLaReina</i>			-1,19 E-01	0,387
<i>DummyLasCondes</i>	-7,04 E-01	0,000	-8,39 E-01	0,000
<i>DummyLoEspejo</i>			-4,04 E-01	0,034
<i>DummyLoPrado</i>			-3,99 E-01	0,030
<i>DummyMacul</i>	-4,65 E-01	0,002	-6,10 E-01	0,000
<i>DummyMaipu</i>			-2,84 E-01	0,006
<i>DummyÑuñoa</i>			-2,03 E-01	0,094
<i>DummyPeñalolen</i>	-5,07 E-01	0,000	-6,58 E-01	0,000
<i>DummyProvidencia</i>	-7,46 E-01	0,000	-9,03 E-01	0,000
<i>DummyPudahuel</i>			-2,95 E-01	0,025
<i>DummyRecoleta</i>	-7,12 E-01	0,000	-8,59 E-01	0,000
<i>DummyRenca</i>			-2,86 E-01	0,065
<i>DummySanJoaquin</i>	-3,09 E-01	0,031	-4,55 E-01	0,002
<i>DummySanMiguel</i>	-9,31 E-01	0,000	-1,11 E+00	0,000
<i>DummyVitacura</i>			-6,87 E-02	0,662
<i>HABITACIONAL</i>	1,73 E-05	0,030	1,71 E-05	0,032
<i>OFICINA</i>	-1,08 E-04	0,062	-1,25 E-04	0,024
<i>SALUD</i>			3,38 E-03	0,193
<i>ln_fv</i>	3,85 E-01	0,000	3,74 E-01	0,000
<i>ln_fp</i>	7,20 E-02	0,000	7,80 E-02	0,000
<i>ln_fbTS</i>	3,04 E-01	0,000	3,20 E-01	0,000
<i>ln_transbordos</i>	3,33 E-02	0,001	3,34 E-02	0,001
<i>Sobredispersión</i>	2,66 E-01		2,48 E-01	
<i>Desvianza</i>	812,0		816,4	
<i>Chi-cuadrado de Pearson</i>	810,3		797,6	
<i>Logaritmo de verosimilitud</i>	-2664,3		-2646,0	
<i>Criterio de información Akaike (AIC)</i>	5368,6		5360,0	

Tabla 6-16: Resultados de la modelación Binomial Negativa para accidentes tipo no atropello.

Fuente: Elaboración propia.

En la Tabla 6-16 se observan variables con estimadores negativos para todas las variables dummy incluidas, además de la variable oficina.

Por otro lado, la significancia de las variables es superior al 90% excepto en los casos de dummy La Cisterna, dummy La Granja, dummy La Reina, dummy Vitacura y el uso de suelo salud.

Parámetros	Acc atropello 1		Acc atropello 2	
	B	Sig.	B	Sig.
<i>(Intersección)</i>	-5,12 E+00	0,000	-5,37 E+00	0,000
<i>DummyCerrillos</i>			-1,32 E+00	0,001
<i>DummyLaFlorida</i>	-4,75 E-01	0,002	-5,14 E-01	0,001
<i>DummyLasCondes</i>	-6,02 E-01	0,000	-5,15 E-01	0,004
<i>DummyLoPrado</i>			-8,54 E-02	0,730
<i>DummyMacul</i>	-4,79 E-01	0,041	-4,74 E-01	0,039
<i>DummyMaipu</i>	-6,81 E-01	0,000	-7,15 E-01	0,000
<i>DummyPeñalolen</i>	-4,19 E-01	0,035	-3,80 E-01	0,050
<i>DummyRecoleta</i>			-2,47 E-01	0,208
<i>DummySanJoaquin</i>	-5,76 E-01	0,022	-5,79 E-01	0,020
<i>DummySanMiguel</i>	-1,69 E+00	0,000	-1,69 E+00	0,000
<i>Vialidadkmsinautopistas</i>	9,73 E-03	0,000	1,05 E-02	0,000
<i>COMERCIO</i>	3,87 E-04	0,007	4,12 E-04	0,004
<i>INDUSTRIA</i>	-2,92 E-03	0,000	-2,35 E-03	0,003
<i>CULTO</i>	1,39 E-02	0,005	1,17 E-02	0,016
<i>SALUD</i>			2,75 E-03	0,405
<i>SumadeAvalúoFiscaldelospredios</i>			-1,42 E-13	0,236
<i>ln_fv</i>	1,79 E-01	0,000	2,25 E-01	0,000
<i>ln_fp</i>	1,17 E-01	0,000	1,43 E-01	0,000
<i>ln_fm</i>			-4,55 E-02	0,004
<i>ln_fbTS</i>	4,24 E-01	0,000	4,11 E-01	0,000
<i>ln_transbordos</i>	5,61 E-02	0,000	5,03 E-02	0,001
<i>Sobredispersión</i>	3,15 E-01		2,86 E-01	
<i>Desvianza</i>	806,6		804,7	
<i>Chi-cuadrado de Pearson</i>	791,1		776,3	
<i>Logaritmo de verosimilitud</i>	-1431,9		-1417,8	
<i>Criterio de información Akaike (AIC)</i>	2897,8		2881,6	

Tabla 6-17: Resultados de la modelación Binomial Negativa para accidentes tipo atropello.
Fuente: Elaboración propia.

Para el modelo de accidentes tipo atropello en la Tabla 6-17 todas las variables dummy poseen coeficientes estimados negativos, además de la variable industria, suma de avalúo fiscal de predios y el flujo de motocicletas.

Las variables no significativas presentadas en los modelos 2 son la dummy de Lo Prado, dummy Recoleta, uso de suelo salud y suma de avalúo fiscal de predios.

La sobredispersión, al igual que los 2 modelos anteriores, presentan diferencias de alrededor de 2 a 3 centésimas.

Parámetros	Lesionados 1		Lesionados 2	
	B	Sig.	B	Sig.
(Intersección)	-3,81 E+00	0,000	-3,62 E+00	0,000
DummyCerrillos			-1,55 E+00	0,000
DummyElBosque	7,02 E-01	0,000	6,15 E-01	0,000
DummyLaGranja	1,24 E+00	0,000	1,18 E+00	0,000
DummyLaPintana	8,04 E-01	0,000	7,12 E-01	0,000
DummyLaReina	3,91 E-01	0,012	2,88 E-01	0,059
DummyMaipu			-2,40 E-01	0,043
DummyÑuñoa	7,13 E-01	0,000	6,29 E-01	0,000
DummyPAC	6,51 E-01	0,000	6,08 E-01	0,000
DummyQtaNormal	1,04 E+00	0,000	9,99 E-01	0,000
DummyRecoleta			-4,01 E-01	0,011
DummySanMiguel	-1,44 E+00	0,000	-1,48 E+00	0,000
DummySanRamon			-5,40 E-01	0,015
DummyVitacura	5,67 E-01	0,001	5,09 E-01	0,003
DummyPuenteAlto	4,45 E-01	0,000	3,35 E-01	0,004
HABITACIONAL			2,59 E-05	0,005
ln_fv	3,42 E-01	0,000	3,21 E-01	0,000
ln_fp	6,76 E-02	0,001	5,30 E-02	0,014
ln_fbTS	3,11 E-01	0,000	3,33 E-01	0,000
ln_transbordos	3,88 E-02	0,001	3,40 E-02	0,002
Sobredispersión	3,80 E-01		3,43 E-01	
Desvianza	838,5		830,2	
Chi-cuadrado de Pearson	787,9		805,3	
Logaritmo de verosimilitud	-2482,4		-2449,6	
Criterio de información Akaike (AIC)	4996,7		4941,1	

Tabla 6-18: Resultados de la modelación Binomial Negativa para lesionados.

Fuente: Elaboración propia.

En los modelos para lesionados, el resultado mostrado en la Tabla 6-18 para la modelación Binomial Negativa, los estimadores de las variables dummy son positivos, excepto para el caso de Cerrillos, Maipú, Recoleta, San Miguel y San Ramón.

Todas las variables resultan significativas para esta formulación.

Parámetros	Fallecidos 1		Fallecidos 2	
	B	Sig.	B	Sig.
<i>(Intersección)</i>	-5,84 E+00	0,000	-6,29 E+00	0,000
<i>DummyLaPintana</i>	9,80 E-01	0,007	1,14 E+00	0,003
<i>DummyQtaNormal</i>			5,62 E-01	0,156
<i>DummyRenca</i>	7,21 E-01	0,075	7,84 E-01	0,052
<i>DummySanJoaquin</i>	7,38 E-01	0,047	7,98 E-01	0,031
<i>DummyPuenteAlto</i>	6,65 E-01	0,014	7,42 E-01	0,006
<i>HABITACIONAL</i>	4,05 E-05	0,057	3,86 E-05	0,073
<i>ln_fv</i>			1,09 E-01	0,196
<i>ln_fbTS</i>	6,38 E-01	0,000	5,50 E-01	0,000
<i>Sobredispersión</i>	6,59 E-01		6,29 E-01	
<i>Desvianza</i>	530,2		531,3	
<i>Chi-cuadrado de Pearson</i>	743,9		741,0	
<i>Logaritmo de verosimilitud</i>	-517,5		-515,9	
<i>Criterio de información Akaike (AIC)</i>	1051,0		1051,8	

Tabla 6-19: Resultados de la modelación Binomial Negativa para fallecidos.

Fuente: Elaboración propia.

Todas las variables de la Tabla 6-19 poseen coeficientes positivos.

Las variables no significativas corresponden a la dummy Quinta Normal y al logaritmo natural del flujo de automóviles.

La sobredispersión difiere entre los modelos tipo 1 y 2 por 3 centésimas.

6.6 Análisis de las variables en los modelos

En este apartado se realizará un análisis de los signos de los estimadores y de su participación en la explicación de la siniestralidad según los modelos calculados.

Primero, se puede observar que algunas variables, por su correlación y significancia en los modelos, no eran relevantes para incluirlas en los modelos. Estas variables son:

- Dummy Huechuraba.
- Dummy Lo Barnechea.
- Dummy San Bernardo.
- Área de la zona en km².
- Vialidad total incluyendo autopistas.
- Total de predios.
- Deporte y recreación.
- Educación y cultura.
- Hotel / Motel.
- Bienes comunes.
- Bodega y almacenaje.
- Minería.
- Administración pública y defensa.
- Transporte y telecomunicaciones.
- Otros no considerados.
- Sitio eriazo.
- Estacionamiento.
- Veh-km.
- Flujo de bicicletas

Todas las variables anteriores tienen algún nivel de correlación con las variables incluidas, y, por lo tanto, es lógico que su significancia se vea afectada.

Las variables dummies tienen por objetivo identificar efectos no corregidos por las variables explicativas cuantitativas utilizadas, en donde solo por pertenecer a una comuna ya se pueden identificar características diversas como, por ejemplo, estado de la infraestructura vial, composición de la población, motorización, clima, entre otras. El efecto que tienen estas variables puede ser positivo o negativo, indicando que la pertenencia de una zona a una comuna puede influir en si disminuyen los riesgos de siniestralidad o aumentan.

Un resumen de los signos de las variables incluidas en la modelación se presenta en la Tabla 6-20.

Variable	Signo del estimador
(Intersección)	Negativo
DummyStgo	Negativo
DummyCerrillos	Negativo
DummyCerroNavia	Negativo
DummyConchali	Negativo
DummyElBosque	Acc. totales 2 y acc. no atropello 2 signo negativo, lesionados 1 y 2 signo positivo
DummyEstacionCentral	Negativo
DummyIndependencia	Negativo
DummyLaCisterna	Negativo
DummyLaFlorida	Negativo
DummyLaGranja	Acc. no atropello 2 signo negativo, lesionados 1 y 2 signo positivo
DummyLaPintana	Positivo
DummyLaReina	Acc. no atropello 2 signo negativo, lesionados 1 y 2 signo positivo
DummyLasCondes	Negativo
DummyLoEspejo	Negativo
DummyLoPrado	Negativo
DummyMacul	Negativo
DummyMaipu	Negativo
DummyÑuñoa	Acc. no atropello 2 signo negativo, lesionados 1 y 2 signo positivo
DummyPAC	Positivo
DummyPeñalolen	Negativo
DummyProvidencia	Negativo
DummyPudahuel	Negativo
DummyQtaNormal	Positivo
DummyRecoleta	Negativo
DummyRenca	Acc. totales 2 y acc. no atropello 2 signo negativo, fallecidos 1 y 2 signo positivo
DummySanJoaquin	Acc. totales, acc. no atropello y acc. atropello, 1 y 2, signo negativo, fallecidos signo positivo
DummySanMiguel	Negativo
DummySanRamon	Acc. totales 2 signo positivo, lesionados 2 signo negativo
DummyVitacura	Acc. no atropello 2 signo negativo, lesionados 1 y 2 signo positivo
DummyPuenteAlto	Positivo
Vialidadkmsinautopistas	Positivo
Habitacional	Positivo
Oficina	Negativo
Comercio	Positivo
Industria	Negativo
Culto	Positivo
Salud	Positivo
SumadeAvalúoFiscaldelospredios	Negativo
In_fv	Positivo
In_fp	Positivo
In_fm	Negativo
In_fbTS	Positivo
In_transbordos	Positivo

Tabla 6-20: Signos de los estimadores para las variables independientes utilizadas.

Fuente: Elaboración propia.

Que el signo del estimador sea negativo implica que la variable, al aumentar su valor, va a disminuir la ocurrencia de accidentes, atropellos, no atropellos, cantidad de lesionados o la cantidad de fallecidos. Por el contrario, si el estimador es positivo, quiere decir que el aumento de valor de la variable va a generar un aumento en la cantidad de accidentes. Cabe recordar que las variables dummy solo poseen valor 1 si pertenece a la comuna, o 0 si no.

Se puede observar de la Tabla 6-20 que para la mayoría de las variables dummy de pertenencia de una zona a una comuna, sus estimadores son negativos. Cambian los signos en algunos modelos, como para la comuna de El Bosque, La Granja, La Reina, Ñuñoa, Renca, San Joaquín, San Ramón y Vitacura.

Las comunas donde el estimador resulta ser positivo son: La Pintana, Pedro Aguirre Cerda, Quinta Normal y Puente Alto. Esto quiere decir que, si una zona pertenece a esas comunas, va a aumentar la siniestralidad de tránsito. Se destaca el hecho de que estas variables son utilizadas en distintos modelos, y no necesariamente en todos. Un resumen de las variables que se utilizan en la modelación se presenta en la Tabla 6-21.

Sobre los signos de los estimadores de las variables que no son dummy, los positivos son: Vialidad sin autopistas, Habitacional, Comercio, Culto, Salud, Flujo vehicular, Flujo peatonal, Flujo de buses del Transantiago y los Transbordos. Los estimadores negativos son: Oficina, Industria, Suma de avalúo fiscal de predios y el flujo de motocicletas.

De lo observado anteriormente, que el estimador de la variable habitacional sea positivo significa que, a mayor cantidad de predios con destino de uso de suelo final destinado a la habitación, aumentará la tasa de accidentabilidad, lo cual es esperable. Lo mismo sucede para la variable de comercio, culto y salud, que son usos de suelo que atraen y generan muchos viajes, de manera sectorizada, aumentando así la exposición de las personas a los riesgos de tránsito.

Los flujos de automóviles, peatones y buses del Transantiago aumentan la tasa de accidentes si es que existe un aumento del valor de estas variables. Esto también es esperable por el mismo fenómeno: aumenta la exposición de las personas a los accidentes de tránsito. Lo contrario sucede con la variable de flujo de motocicletas, que se utiliza solo para el modelo de accidentes atropello, tipo 2. Esto último se podría explicar debido a que el flujo de motocicletas es relativamente bajo (1.117 viajes promedio por día por zona) y puede existir una correlación que no se observe en los estadísticos ni en los test.

Variable independiente	Acc. totales 1	Acc. totales 2	Acc. No atropello 1	Acc. No atropello 2	Acc. atropello 1	Acc. atropello 2	Lesionados 1	Lesionados 2	Fallecidos 1	Fallecidos 2
DummyStgo	X	X	X	X						
DummyCerrillos	X	X	X	X		X		X		
DummyCerroNavia		X		X						
DummyConchali	X	X	X	X						
DummyEIBosque		X		X			X	X		
DummyEstacionCentral	X	X	X	X						
DummyIndependencia	X	X	X	X						
DummyLaCisterna		X		X						
DummyLaFlorida		X		X	X	X				
DummyLaGranja				X			X	X		
DummyLaPintana							X	X	X	X
DummyLaReina				X			X	X		
DummyLasCondes	X	X	X	X	X	X				
DummyLoEspejo		X		X						
DummyLoPrado		X		X		X				
DummyMacul	X	X	X	X	X	X				
DummyMaipu	X	X		X	X	X		X		
DummyÑuñoa				X			X	X		
DummyPAC							X	X		
DummyPeñalolen	X	X	X	X	X	X				
DummyProvidencia	X	X	X	X						
DummyPudahuel		X		X						
DummyQtaNormal							X	X		X
DummyRecoleta	X	X	X	X		X		X		
DummyRenca		X		X					X	X
DummySanJoaquin	X	X	X	X	X	X			X	X
DummySanMiguel	X	X	X	X	X	X	X	X		
DummySanRamon		X						X		
DummyVitacura				X			X	X		
DummyPuenteAlto							X	X	X	X
Vialidadkmsinautopistas					X	X				
HABITACIONAL	X	X	X	X				X	X	X
OFICINA		X	X	X						
COMERCIO					X	X				
INDUSTRIA					X	X				
CULTO					X	X				
SALUD		X		X		X				
SumadeAvalúoFiscaldelospredios						X				
In_fv	X	X	X	X	X	X	X	X		X
In_fp	X	X	X	X	X	X	X	X		
In_fm						X				
In_fbTS	X	X	X	X	X	X	X	X	X	X
In_transbordos	X	X	X	X	X	X	X	X		

Tabla 6-21: Resumen de las variables explicativas que se utilizan en cada modelo.
Fuente: Elaboración propia.

Los coeficientes en el caso de los modelos de accidentes totales son todos negativos para las dummies de pertenencia a alguna comuna, excepto para la comuna de San Ramón. Algunas diferencias sustantivas incluso mayores a 1 orden de magnitud. Esto indica que algunas comunas tienen mayor prevalencia de accidentes que otras, y que existen factores exógenos que solo están siendo percibidos por esta variable.

Las variables dummy para los modelos de no atropello y atropellos también poseen estimadores negativos en todos los casos. La situación cambia cuando se observan los estimadores para los modelos de cantidad de lesionados y fallecidos. En estos casos los estimadores son generalmente positivos, excepto para las variables de: Maipú, Recoleta, San Miguel y San Ramón. En los modelos de fallecidos son todas las variables dummy positivas. Estos fenómenos se explican tal cual se mencionó anteriormente: existen factores exógenos a las variables de la modelación que si son percibidos por estas variables y no por las demás.

El caso de la variable de vialidad sin autopistas, incluida en los modelos de atropellos presenta un valor positivo para su coeficiente indicando que, a mayor infraestructura de caminos por zona, mayor será la frecuencia de atropellos. Esto también es esperable debido a que mayor vialidad representa mayor exposición al tránsito.

En el caso de las variables de comercio y oficina, que presentan una alta correlación, se puede notar que en los modelos de atropellos se incluye comercio en vez de oficina y que posee un coeficiente positivo indicando que, a mayor cantidad de predios con este destino de uso de suelo, mayor será la cantidad de atropellos. Dicha situación se condice con la realidad. Sobre la variable oficina, en los modelos donde se incluye, siempre tiene coeficiente negativo, lo que sugiere analizar el porqué de esta situación. Se puede deber a que las oficinas tienden a generar y atraer viajes en horarios punta, donde el riesgo por exposición es mayor, pero las consecuencias son menores en los accidentes de tránsito debido al uso más eficiente del espacio vial en transporte público y modos alternativos a este, así como la reducción de la velocidad de flujo a nivel general por congestión y por consiguiente la reducción en las consecuencias en la salud de los siniestrados.

La variable habitacional presenta siempre coeficiente positivo, lo cual plantea una conclusión lógica esperable. La cantidad de hogares presentes en una zona son capaces de generar y atraer viajes, por lo que son fuente importante del aumento del nivel de flujo y por lo tanto de la exposición de los usuarios. Era de esperar que, a mayor densidad habitacional, mayor sea la cantidad de accidentes, lo que se refleja en el signo del coeficiente.

El caso de la industria es similar al de comercio. Su coeficiente es negativo en los modelos de atropello implicando que, a mayor cantidad de industrias, menor la cantidad de atropellos. Esto se puede deber a un efecto parecido al de la oficina, en donde el nivel de flujo y exposición de los usuarios está temporalmente segregado y afecto a las condiciones laborales de la industria. Además, generalmente las industrias tienden a tener predios de mayor tamaño que otros y, por lo tanto, a nivel de uso de suelo, es menos concurrido que otro tipo.

Sobre los predios con destino culto y salud, ambos presentan siempre coeficientes positivos siendo incluidos en los modelos de accidentes totales, atropellos y no atropellos. Presentan correlación con la variable habitacional, por lo que pueden estar sesgando el efecto de esta última variable en los modelos.

La variable de suma de los avalúos fiscales de los predios está presente en los modelos de atropello tipo 2, y su estimador es negativo. Esto indica que, a mayor avalúo de los predios, menor será la tasa de accidentes del tipo atropello.

Los coeficientes positivos de las variables de flujo indican que estas son importantes en el diagnóstico de la seguridad vial de una ciudad. Naturalmente que el flujo y sus confluencias son la raíz principal de los efectos nocivos de la siniestralidad.

El fenómeno de que una persona en cualquier modo es un potencial peatón se intenta abordar incluyendo la variable de transbordos en la modelación. Esta resulta significativa en todos los modelos, excepto en los de fallecidos, y su coeficiente es siempre positivo, indicando que, a mayor cantidad de transbordos, mayor es la exposición de los peatones y por lo tanto mayor el riesgo de ocurrencia de un amargo episodio vial.

Sobre la participación en la tasa de accidentes de las personas, se puede notar que los coeficientes del flujo vehicular son mayores que los del flujo de buses, y a su vez estos son mayores que los coeficientes para los peatones, en los modelos de accidentes no atropello. Para los modelos de accidentes totales y cantidad de lesionados, los estimadores del flujo de buses son mayores que los de automóviles, y estos mayores que para los peatones, aunque el estimador del flujo peatonal es de mayor valor que para los otros modelos. Esto último es esperable debido a que los peatones son los principales involucrados en los accidentes atropellos y, por lo tanto, un aumento de la exposición de estos aumenta el riesgo de siniestralidad, al mismo tiempo que aumentan las consecuencias en la salud de los accidentados.

Por otro lado, el coeficiente de los buses, para los modelos de atropellos es mayor que el del flujo de automóviles, y estos mayores que el estimador del flujo peatonal. Esto puede ser debido a que la variable de flujo peatonal solo considera una proporción de los peatones. Un viajero que, por ejemplo, cambia de modo en su viaje, es un peatón en una fracción de su viaje, pero que de igual forma aumenta el riesgo debido a su exposición con el tránsito. Este fenómeno puede no estar bien percibido por las variables ya que en la encuesta origen destino no se declaran estas etapas intermedias ni menos se puede percibir el efecto real ya que no se cuentan con mediciones de flujo peatonal certeras para todas las zonas de la ciudad.

Sobre los modelos de cantidad de fallecidos, solo resultan significativas las variables de flujo de buses y el flujo vehicular. Esto puede ser debido al fenómeno antes explicado de que las variables perciben efectos exógenos o pueden no ser representativos de manera absoluta del riesgo que presenta un viajero en dicho modo.

6.7 Análisis de efecto marginal sobre las tasas de accidentes de una persona más en modo auto, bus o caminata

Para analizar el efecto marginal de una persona viajando en vehículo, bus o como peatón, se considera el estudio de mediciones de tráfico (SECTRA, 2013) en donde se calcularon los aforos y perfiles de carga de servicios en el Gran Santiago. Este considera 406 puntos de medición esparcidos en la ciudad. Las tasas de ocupación allí reportadas y presentadas en el apartado 5.6 muestran entre 21 y 39 pasajeros por tipo de bus (30,88 promedio ponderado) y 1,45 personas por automóvil. Debido a que la variable de flujo de buses de la modelación fue obtenida de los programas de operación del Transantiago, las unidades de esta son buses promedio por día, por zona. Es decir, no tienen incluido el efecto de la cantidad de viajes que se generan por persona. Al mismo tiempo, el flujo vehicular considera cantidad de vehículos por zona, que no representa a la cantidad de personas que viajan.

Luego, si se considera que en bus viajan en promedio 30,88 personas y en un automóvil la tasa de ocupación es de 1,45 personas/vehículo, y además se nota que el estimador en el caso de los flujos de tránsito representan la elasticidad de dicha variable frente a la tasa de accidentalidad, como se puede notar en la formulación de la Ecuación 13, el efecto que tendría una persona viajando en un autobús debiese ser menor que el efecto de una persona viajando en automóvil, debido a que cada bus transporta 21,3 veces más pasajeros. Lo anterior depende totalmente de la cantidad de viajes en cada modo que se estén generando en cada zona.

La elasticidad representa la variación porcentual de una variable al variar en un uno por ciento la otra variable de estudio. Es decir, si observamos el estimador del flujo vehicular para el modelo binomial negativo tipo 1 de los accidentes totales, este es de valor 0,356. Es decir, cuando aumente en un uno por ciento el nivel de flujo de automóviles, debiese variar en un 0,356% la tasa de accidentes totales.

En cada automóvil viajan 1,45 personas, por lo tanto, el aumento marginal de una persona representa el inverso multiplicativo de ese valor, correspondiente a 0,69 vehículos por persona. El caso de los buses es análogo, y el recíproco de la tasa de ocupación vale 0,03 buses que corresponden a una persona. Con los valores de los estimadores se puede obtener cuánto varía porcentualmente la tasa de accidentes cuando varía en 0,69 el nivel de flujo de automóviles y cuánto varía la tasa de accidentes cuando varía en 0,03 el nivel de flujo de buses. Este aumento es, simplemente, la multiplicación del estimador (o elasticidad) por el valor del aumento marginal representativo a una persona viajando en automóvil o en bus para cada zona. Es decir, es necesario observar qué porcentajes representan estos aumentos en cada una de las zonas, en cada una de las variables observadas. Para el caso de los peatones, el aumento marginal porcentual en la tasa de accidentes se obtiene como el recíproco del flujo peatonal de cada zona.

En base a lo anterior, se calcula para cada zona este aumento porcentual y se multiplica por la elasticidad (estimador) de cada modelo.

En la Tabla 6-22, Tabla 6-23, Tabla 6-24 y Tabla 6-25 se muestra la cantidad de zonas donde cada variable tuvo aumento porcentual en la tasa de accidentes mayor, intermedio o menor que el resto de las variables. Es decir, la columna con valor 1 quiere decir que la variable “fv”, “fp” o “fbTS” obtuvo el mayor aumento porcentual en esa cantidad de zonas. La columna con valor 2 obtuvo el aumento porcentual intermedio al respecto de las otras variables, y análogamente, la columna con valor 3 corresponde a la cantidad de zonas en que el aumento porcentual en la tasa de accidentes fue menor. Se presenta cada tabla según cada variable explicada, y luego, se concluye sobre lo que allí se observa. Además, cabe recordar que, según la nomenclatura utilizada en este trabajo, “fv” corresponde al flujo de automóviles, “fp” al flujo peatonal y “fbTS” al flujo de buses del Transantiago.

Tipo de modelo	Formulación	Variable independiente de flujo	1	2	3
Acc. totales 1	Normal	fv	402	152	160
		fp	219	217	278
		fbTS	93	345	276
	Poisson	fv	499	175	40
		fp	41	38	635
		fbTS	174	501	39
	Binomial Negativa	fv	548	159	7
		fp	5	6	703
		fbTS	161	549	4
Acc. totales 2	Normal	fv	501	211	2
		fp	1	1	712
		fbTS	212	502	0
	Poisson	fv	461	252	1
		fp	0	1	713
		fbTS	253	461	0
	Binomial Negativa	fv	464	250	0
		fp	0	0	714
		fbTS	250	464	0

Tabla 6-22: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de accidentes totales.

Fuente: Elaboración propia.

De la Tabla 6-22 se puede notar, a modo de ejemplo, que la variable de flujo de automóviles (fv) para el modelo de accidentes totales tipo 1, en la formulación Normal, en 402 zonas fue el aumento mayor en la tasa de accidentabilidad en términos porcentuales, en 152 el aumento intermedio y en 160 el que representó el menor aumento porcentual de los accidentes totales. Se destacan los valores mayores para cada formulación.

Por lo tanto, se puede notar que, en todas las formulaciones para los dos tipos de modelos de accidentes totales, se repite el mismo patrón: el aumento en una persona viajando en automóvil representa el mayor aumento porcentual en la tasa de accidentes, luego, será menor el aumento porcentual en la tasa para una persona en buses del Transantiago, y finalmente, el aumento en un peatón representará el menor aumento de

la tasa de accidentes. Por lo tanto, se puede concluir que, en los modelos de accidentes totales, para la mayoría de las zonas, una persona más viajando en automóvil aumenta en mayor medida la tasa de accidentes totales que una persona más viajando en transporte público, y a su vez, el peatón representa el menor aumento porcentual en la siniestralidad.

Tipo de modelo	Formulación	Variable independiente de flujo	1	2	3
Acc. No atropello 1	Normal	fv	494	220	0
		fp	0	0	714
		fbTS	220	494	0
	Poisson	fv	526	188	0
		fp	0	0	714
		fbTS	188	526	0
	Binomial Negativa	fv	578	136	0
		fp	0	0	714
		fbTS	136	578	0
Acc. No atropello 2	Normal	fv	600	114	0
		fp	0	0	714
		fbTS	114	600	0
	Poisson	fv	613	101	0
		fp	0	0	714
		fbTS	101	613	0
	Binomial Negativa	fv	635	79	0
		fp	0	0	714
		fbTS	79	635	0

Tabla 6-23: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de accidentes no atropello.
Fuente: Elaboración propia.

Para los modelos de accidentes no atropello de la Tabla 6-23 se puede notar la misma conclusión que para los modelos de accidentes totales: en la mayoría de las zonas, una persona más viajando en un automóvil representa el mayor aumento porcentual de la tasa de accidentes no atropello. El efecto de una persona más viajando en un bus del Transantiago será menor que en automóvil, y finalmente, el aumento porcentual de un peatón más en cada zona implicará el menor aumento porcentual en la tasa de accidentes no atropello.

Tipo de modelo	Formulación	Variable independiente de flujo	1	2	3
Acc. atropello 1	Normal	fv	520	194	0
		fp	0	0	714
		fbTS	194	520	0
	Poisson	fv	280	434	0
		fp	0	0	714
		fbTS	434	280	0
	Binomial Negativa	fv	84	630	0
		fp	0	0	714
		fbTS	630	84	0
Acc. atropello 2	Normal	fv	60	654	0
		fp	0	0	714
		fbTS	654	60	0
	Poisson	fv	22	692	0
		fp	0	0	714
		fbTS	654	22	0
	Binomial Negativa	fv	6	708	0
		fp	0	0	714
		fbTS	708	6	0

Tabla 6-24: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de accidentes atropello.

Fuente: Elaboración propia.

En la Tabla 6-24 se rompe la tendencia de los modelos de accidentes totales y accidentes no atropello. En la mayoría de las formulaciones, excepto para los modelos Normal de accidentes atropello tipo 1, el aumento en una persona más viajando en bus de Transantiago tiene un mayor aumento porcentual en la tasa de accidentes atropello. Luego, el efecto de una persona más viajando en auto será menor, y certeramente en todas las zonas el efecto de un peatón más será el menor aumento porcentual en la siniestralidad.

Por lo tanto, de manera general en casi todas las zonas, el efecto marginal sobre el aumento de los accidentes atropello de una persona más viajando en bus será mayor que el efecto marginal de un automovilista, y a su vez el efecto marginal de un peatón será mucho menor que estas dos variables de flujo.

Tipo de modelo	Formulación	Variable independiente de flujo	1	2	3
Lesionados 1	Normal	fv	6	708	0
		fp	0	0	714
		fbTS	708	6	0
	Poisson	fv	6	708	0
		fp	0	0	714
		fbTS	708	6	0
	Binomial Negativa	fv	7	707	0
		fp	0	0	714
		fbTS	707	7	0
Lesionados 2	Normal	fv	6	708	0
		fp	0	0	714
		fbTS	708	6	0
	Poisson	fv	6	708	0
		fp	0	0	714
		fbTS	708	6	0
	Binomial Negativa	fv	6	708	0
		fp	0	0	714
		fbTS	708	6	0

Tabla 6-25: Cantidad de zonas en las que cada variable independiente de flujos produjo un aumento porcentual mayor (1), intermedio (2) o menor (3) en las tasas de lesionados.

Fuente: Elaboración propia.

Para los modelos de lesionados de la Tabla 6-25 sucede casi lo mismo que en los modelos de accidentes atropello: el efecto marginal mayor sobre el aumento de la tasa de lesionados será de una persona más viajando en bus del Transantiago, luego vendrá el efecto de un automovilista, y finalmente un peatón aumenta las tasas en menor proporción. Se observa también que, en todas las zonas, el efecto de un peatón más viajando será menor que en los otros dos modos en el análisis. En solo 6 o 7 zonas se invertirán los efectos marginales para el flujo de buses y flujo vehicular.

Cabe destacar sobre todo el análisis de las últimas cuatro tablas, que siempre los estimadores de las variables de flujo vehicular, peatonal y de buses son positivos, y, por lo tanto, siempre se habla de un aumento porcentual debido a que entonces la elasticidad es positiva y un aumento porcentual de flujo se condice con un aumento en la tasa de accidentes.

No se consideraron los modelos de fallecidos debido a que no se incluyen las tres variables de flujo que se discuten en esta sección. Además, en algunas zonas había flujos cero, por lo tanto, el análisis era inválido. En particular, son 22 zonas que no poseen alguno de los tres flujos. De esto es que solo habrá 714 zonas en las tablas anteriores, y no 736 como se consideran en la modelación original.

En esta modelación, las variables de flujo vehicular y flujo de buses no consideran las diversas etapas de un viaje que una persona puede realmente enfrentar, aumentando el nivel de exposición de los usuarios peatones. Es decir, probablemente una persona en vehículo y una persona en transporte público se convierten en peatones en algún momento del viaje. Luego, se podría estar sobre estimando el efecto de los peatones

sobre la generación de accidentes también debido a que se consideran en las formulaciones la cantidad de transbordos, que es una medida de representación del nivel de flujo peatonal en una zona y, por lo tanto, se sobreestima el efecto de la variable peatonal o se lleva parte de ese fenómeno el coeficiente estimado de transbordos. Por otro lado, las tasas de ocupación utilizadas son constantes para toda la ciudad, lo cual también implica un sesgo en los cálculos.

Sobre el resultado en los modelos de atropellos y cantidad de lesionados, se puede notar que la cantidad de kilómetros que los autos recorren al día en Santiago es de 2.145.919.340 veh-km (Camus, 2017) en 1.358.729 vehículos particulares al año 2013 en las 32 comunas de la provincia de Santiago más Puente Alto y San Bernardo (Instituto Nacional de Estadística, 2013). La cantidad de kilómetros recorridos por los buses es de 464 millones al año en 6.493 buses para el año 2013 (Directorio de Transporte Público Metropolitano, 2013). Por lo tanto, los vehículos particulares motorizados producen aproximadamente 1.579 km y los buses 196 km al día. Son 8,1 veces más kilómetros recorridos por vehículos particulares en relación con los buses, lo cual indica su relevancia en la exposición de los vehículos particulares por sobre los buses del transporte público.

Además, para el año 2012 (Carabineros de Chile, 2017), la cantidad de vehículos particulares involucrados en accidentes tipo atropello es de 1.793 y la cifra de buses envueltos es de 403. Para los accidentes no atropello son 25.080 vehículos y 2.915 buses implicados. Esto indica que hay 4,45 veces más autos siniestrados que buses en atropellos, y 8,6 veces más autos dañados que buses en accidentes de tipo no atropellos. Es decir, casi se duplica la proporción de buses involucrados en atropellos. De esto se puede explicar la razón del orden del efecto marginal en atropellos, donde el resultado indica que, una persona más, en general, aumenta el riesgo en transporte público en mayor medida que en vehículo particular y, al mismo tiempo, que caminando. Para la cantidad de lesionados, estas proporciones son similares: 4,46 veces más autos que buses en atropellos, y 7,38 veces para no atropellos. También destaca que hay 8,1 veces más autos que buses involucrados en accidentes totales, y hay 6,81 veces más lesionados en auto que en buses.

6.8 Estimación aplicando el método empírico bayesiano

Dado que el método empírico bayesiano considera un ajuste según el valor de la sobredispersión, solo es aplicable al caso de los modelos Binomial Negativa.

El parámetro de sobredispersión obtenido mediante el método chi-cuadrado de Pearson se presenta en la Tabla 6-26 para los distintos modelos.

Modelo	Sobredispersión
Acc tot-1	0,263
Acc tot-2	0,246
Acc na-1	0,266
Acc na-2	0,248
Acc at-1	0,315
Acc at-2	0,286
Acc les-1	0,380
Acc les-2	0,343
Acc fall-1	0,659
Acc fall-2	0,629

Tabla 6-26: Sobredispersión para los distintos modelos Binomial Negativa.
Fuente: Elaboración propia.

Con estos valores se puede calcular α de la Ecuación 7 para cada zona y posteriormente calcular la estimación para cada zona y para cada variable explicada mediante la fórmula del método empírico bayesiano en la Ecuación 8 considerando el registro histórico de los accidentes. Para la estimación del año 2013 se consideran los accidentes promedio del año 2011 y 2012 debido a que solo se cuenta con dicha información anterior al año de modelación. Para la predicción del capítulo 7 se utiliza el registro histórico como el promedio entre los años 2011 y 2014. Si bien la literatura sugiere considerar más años para los promedios históricos en su uso para la estimación de tasas de accidentes, (Hauer, 1997), la disponibilidad de datos no es tal que permita holgura en este sentido.

6.9 Comparación entre formulaciones

En este apartado se mostrarán las diferencias entre las distintas formulaciones en cuanto a su bondad de ajuste y los estadísticos que indican aquello. También se abordará la comparación de los estimadores y la significancia entre la formulación Normal, Poisson y Binomial Negativa, para cada tipo de variable explicada.

La comparación entre las formulaciones dicta principalmente sobre el nivel de ajuste de los modelos. Esto se puede observar en los coeficientes de la bondad de ajuste devianza, chi-cuadrado de Pearson, log verosimilitud y AIC. En el software SPSS se estiman dichos estadísticos para que mientras pequeños sean, mejor es el ajuste.

La prueba de contraste ómnibus es un test de chi-cuadrado de la razón de verosimilitud del modelo actual frente al modelo de interceptación. Si el valor de significancia es inferior a 0,05, el modelo actual funciona mejor que el modelo nulo. En todos los modelos la prueba de ómnibus entregó una significancia muy cercana a cero, por lo tanto, se puede rechazar la hipótesis nula de que el modelo de interceptación es mejor que el modelo actual, es decir, en todos los modelos se puede considerar que ajustan la realidad en alguna medida significativa.

Modelos	Desviianza			Chi-cuadrado de Pearson			Logaritmo de verosimilitud			AIC		
	Normal	Poisson	Binomial Negativa	Normal	Poisson	Binomial Negativa	Normal	Poisson	Binomial Negativa	Normal	Poisson	Binomial Negativa
Acc totales-1	163606,2	4979,9	811,8	163606,2	5364,8	830,5	-82479,4	-4180,3	-2748,5	164996,9	8398,7	5536,9
Acc totales-2	142122,4	4623,8	815,7	142122,4	4842,9	820,0	-71737,5	-4002,3	-2730,3	143535,1	8064,6	5522,7
Acc no atropello-1	132503,7	4556,5	812,0	132503,7	4873,4	810,3	-66928,2	-3918,2	-2664,3	133894,4	7874,4	5368,6
Acc no atropello-2	112228,2	4199,1	816,4	112228,2	4360,0	797,6	-56790,5	-3739,5	-2646,0	113646,9	7545,0	5360,0
Acc atropello-1	4763,7	1413,8	806,6	4763,7	1447,4	791,1	-3058,2	-1533,2	-1431,9	6148,4	3098,4	2897,8
Acc atropello-2	4252,6	1350,9	804,7	4252,6	1362,9	776,3	-2802,6	-1501,8	-1417,8	5649,3	3047,6	2881,6
Lesionados-1	79414,8	4274,0	838,5	79414,8	4575,9	787,9	-40383,7	-3599,3	-2482,4	80797,4	7228,6	4996,7
Lesionados-2	77440,4	4052,0	830,2	77440,4	4340,6	805,3	-39396,5	-3488,3	-2449,6	78833,1	7016,5	4941,1
Fallecidos-1	291,0	660,4	530,2	291,0	898,3	743,9	-821,9	-523,8	-517,5	1657,7	1061,7	1051,0
Fallecidos-2	289,0	656,3	531,3	289,0	888,7	741,0	-820,8	-521,8	-515,9	1659,7	1061,5	1051,8

Tabla 6-27: Estadísticos de bondad de ajuste de los modelos estimados.
Fuente: Elaboración propia.

Se puede notar de la Tabla 6-27 que, en los modelos de tasa de accidentes totales, accidentes tipo no atropello, tipo atropello y lesionados, siempre todos los estadísticos son menores para la formulación Binomial Negativa. Sin embargo, para el caso de los fallecidos, el modelo Normal tiene mejor Desviianza y chi-cuadrado de Pearson que la formulación Poisson y Binomial Negativa, pero el logaritmo de la verosimilitud y AIC son ligeramente menores para la formulación Binomial Negativa.

Se tienen mejores estadísticos para los modelos tipo 2, excepto para la variable dependiente de cantidad de fallecidos, donde la Desviianza, el logaritmo de la verosimilitud y AIC muestran mejores ajustes para el modelo tipo 1, por menos del 1%.

Es de esperar un mejor resultado de los modelos con la formulación Binomial Negativa puesto que la especificación de Poisson es un caso particular de este cuando la sobredispersión es nula, cosa que se revisó en el apartado 6.4, encontrándose que sí existe sobredispersión en todos los datos de variable explicada. Sobre el caso de la distribución normal, esta tiene por supuesto necesario la normalidad de la variable explicada.

En los gráficos siguientes se muestran los histogramas que prueban que no existe dicha relación normal con los datos y, por lo tanto, no es del todo una formulación correcta que llevará a sesgos importantes en la estimación de los coeficientes de los modelos.

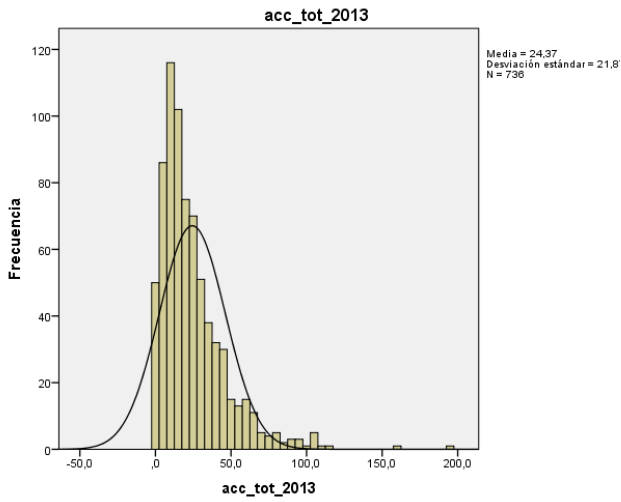


Gráfico 2: Histograma de la cantidad de accidentes totales.
Fuente: Elaboración propia.

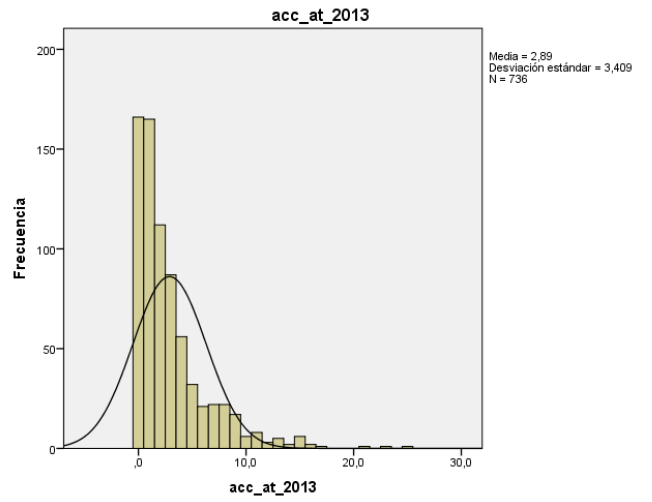


Gráfico 4: Histograma de la cantidad de accidentes tipo atropello.
Fuente: Elaboración propia.

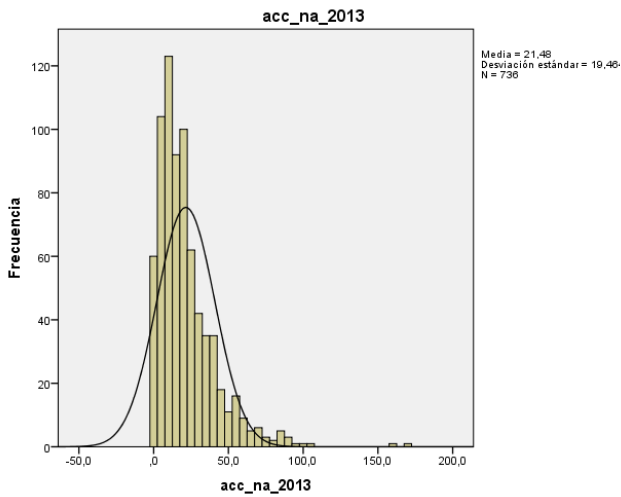


Gráfico 3: Histograma de la cantidad de accidentes tipo no atropello.
Fuente: Elaboración propia.

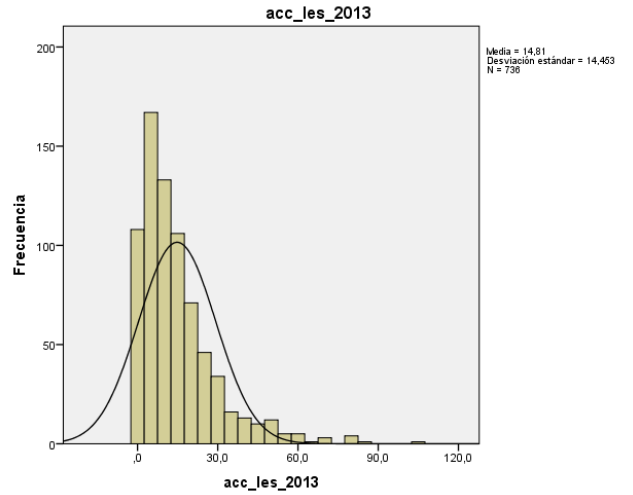


Gráfico 5: Histograma de la cantidad de lesionados.
Fuente: Elaboración propia.

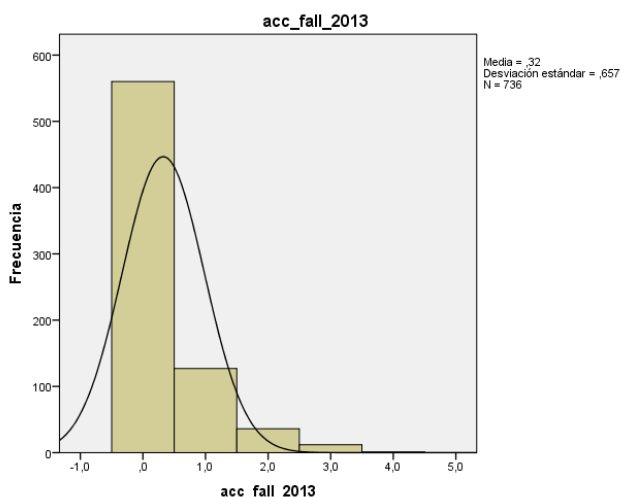


Gráfico 6: Histograma de la cantidad de fallecidos.
Fuente: Elaboración propia.

A continuación, se muestran los resultados de los modelos para cada tipo, según la variable explicada, y luego se comenta sobre las diferencias que se observan entre formulaciones.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-2,956	0,000	-3,198	0,000	-3,170	0,000
DummyStgo	-0,357	0,000	-0,282	0,000	-0,184	0,047
DummyCerrillos	-1,164	0,012	-1,259	0,000	-1,304	0,000
DummyConchali	-0,614	0,011	-0,589	0,001	-0,555	0,000
DummyEstacionCentral	-0,859	0,000	-0,887	0,000	-0,883	0,000
DummyIndependencia	-0,558	0,015	-0,616	0,001	-0,660	0,000
DummyLasCondes	-0,849	0,000	-0,808	0,000	-0,708	0,000
DummyMacul	-0,555	0,001	-0,526	0,000	-0,492	0,001
DummyMaipu	-0,179	0,024	-0,183	0,033	-0,175	0,080
DummyPeñalolen	-0,570	0,001	-0,532	0,000	-0,492	0,000
DummyProvidencia	-0,753	0,000	-0,725	0,000	-0,703	0,000
DummyRecoleta	-0,709	0,000	-0,681	0,000	-0,645	0,000
DummySanJoaquin	-0,386	0,039	-0,370	0,015	-0,350	0,015
DummySanMiguel	-1,227	0,000	-1,118	0,000	-1,019	0,000
HABITACIONAL	1,575E-05	0,004	1,781E-05	0,005	1,982E-05	0,014
In_fv	0,325	0,000	0,342	0,000	0,356	0,000
In_fp	0,077	0,002	0,076	0,000	0,078	0,000
In_fbTS	0,324	0,000	0,339	0,000	0,312	0,000
In_transbordos	0,046	0,000	0,040	0,000	0,036	0,000

Tabla 6-28: Resumen de la modelación para la variable de accidentes totales, tipo 1.
Fuente: Elaboración propia.

La Tabla 6-28 muestra los estimadores y su significancia para los modelos tipo 1 de la variable de accidentes totales. Estos modelos son aquellos en los que todas las variables incluidas, en las tres formulaciones, son significativas a, al menos, un 90%.

Se destaca que ninguno de los coeficientes calculados cambia de signo entre las distintas formulaciones, además de ser relativamente parecidos. En general, la formulación Binomial Negativa tiende a estimar menores valores de los estimadores, en valor absoluto. La mayor diferencia en términos relativos la presenta el estimador de la variable dummy para la comuna de Santiago.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-2,807	0,000	-3,085	0,000	-3,105	0,000
DummyStgo	-0,369	0,000	-0,313	0,000	-0,236	0,013
DummyCerrillos	-1,213	0,006	-1,325	0,000	-1,373	0,000
DummyCerroNavia	-0,215	0,219	-0,246	0,128	-0,278	0,091
DummyConchali	-0,692	0,002	-0,678	0,000	-0,644	0,000
DummyElBosque	-0,390	0,053	-0,434	0,011	-0,460	0,003
DummyEstacionCentral	-0,930	0,000	-0,970	0,000	-0,967	0,000
DummyIndependencia	-0,641	0,003	-0,712	0,000	-0,753	0,000
DummyLaCisterna	-0,258	0,042	-0,212	0,103	-0,138	0,350
DummyLaFlorida	-0,065	0,336	-0,173	0,036	-0,245	0,012
DummyLasCondes	-0,818	0,000	-0,811	0,000	-0,745	0,000
DummyLoEspejo	-0,414	0,117	-0,412	0,046	-0,388	0,040
DummyLoPrado	-0,536	0,007	-0,407	0,019	-0,321	0,076
DummyMacul	-0,612	0,000	-0,596	0,000	-0,566	0,000
DummyMaipu	-0,309	0,000	-0,290	0,001	-0,274	0,006
DummyPeñalolen	-0,664	0,000	-0,621	0,000	-0,579	0,000
DummyProvidencia	-0,906	0,000	-0,782	0,000	-0,732	0,000
DummyPudahuel	-0,364	0,025	-0,310	0,020	-0,248	0,056
DummyRecoleta	-0,775	0,000	-0,764	0,000	-0,727	0,000
DummyRenca	-0,259	0,118	-0,278	0,067	-0,229	0,133
DummySanJoaquin	-0,449	0,011	-0,449	0,002	-0,430	0,003
DummySanMiguel	-1,351	0,000	-1,243	0,000	-1,133	0,000
DummySanRamon	0,247	0,038	0,246	0,076	0,263	0,112
HABITACIONAL	1,009E-05	0,041	1,443E-05	0,018	1,889E-05	0,018
OFICINA	-8,803E-05	0,033	-9,966E-05	0,020	-9,899E-05	0,077
SALUD	0,009	0,000	0,007	0,000	0,004	0,166
In_fv	0,290	0,000	0,312	0,000	0,336	0,000
In_fp	0,085	0,000	0,087	0,000	0,085	0,000
In_fbTS	0,352	0,000	0,361	0,000	0,332	0,000
In_transbordos	0,045	0,000	0,040	0,000	0,038	0,000

Tabla 6-29: Resumen de la modelación para la variable de accidentes totales, tipo 2.

Fuente: Elaboración propia.

El resumen mostrado en la Tabla 6-29 consta de los modelos Normal, Poisson y Binomial Negativo para el caso de la variable explicada de accidentes totales, del tipo 2 en donde las variables incluidas son significativas en al menos una de las tres formulaciones.

Al igual que el caso de los modelos tipo 1 de la Tabla 6-28, no hay diferencias en cuanto al signo de los estimadores. Las diferencias más grandes en cuanto al valor del estimador lo presenta la variable de salud en donde supera el doble de su estimador entre el modelo Binomial Negativo y el modelo Normal. El valor absoluto de los coeficientes de los modelos Binomial Negativo son menores en casi todos los casos.

Sobre las significancias de las variables: La dummy de Cerro Navia es solo significativa para el modelo Binomial Negativo; La dummy de La cisterna es solo significativa para el modelo Normal; Dummy de La florida es no significativa solo para el modelo Normal, al igual que la variable dummy de Lo Espejo; La dummy Renca es solo significativa para el modelo Poisson; Dummy San ramón no es significativa para el modelo Binomial Negativo, al igual que tampoco lo es la variable de salud para dicho modelo.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-3,272	0,000	-3,483	0,000	-3,444	0,000
DummyStgo	-0,335	0,000	-0,263	0,000	-0,176	0,059
DummyCerrillos	-1,139	0,014	-1,229	0,000	-1,276	0,000
DummyConchali	-0,612	0,014	-0,601	0,001	-0,584	0,000
DummyEstacionCentral	-0,998	0,000	-1,026	0,000	-1,029	0,000
DummyIndependencia	-0,585	0,015	-0,639	0,001	-0,684	0,000
DummyLasCondes	-0,816	0,000	-0,778	0,000	-0,704	0,000
DummyMacul	-0,532	0,001	-0,500	0,000	-0,465	0,002
DummyPeñalolen	-0,561	0,002	-0,531	0,000	-0,507	0,000
DummyProvidencia	-0,762	0,000	-0,750	0,000	-0,746	0,000
DummyRecoleta	-0,751	0,000	-0,733	0,000	-0,712	0,000
DummySanJoaquin	-0,332	0,067	-0,322	0,033	-0,309	0,031
DummySanMiguel	-1,152	0,000	-1,033	0,000	-0,931	0,000
HABITACIONAL	1,483E-05	0,008	1,603E-05	0,012	1,731E-05	0,030
OFICINA	-9,782E-05	0,039	0,000	0,026	0,000	0,062
In_fv	0,353	0,000	0,371	0,000	0,385	0,000
In_fp	0,079	0,002	0,073	0,001	0,072	0,000
In_fbTS	0,313	0,000	0,328	0,000	0,304	0,000
In_transbordos	0,043	0,000	0,037	0,000	0,033	0,001

Tabla 6-30: Resumen de la modelación para la variable de accidentes no atropello, tipo 1.

Fuente: Elaboración propia.

La Tabla 6-30 muestra el resumen de la modelación para la variable explicada de accidentes no atropello, para los modelos tipo 1 donde todas las variables son significativas para todos los modelos.

Nuevamente los signos de los estimadores no cambian entre formulaciones y la mayoría de las veces los valores absolutos de los coeficientes calculados con el modelo Binomial Negativo son menores que para Poisson y para Normal. Las mayores diferencias porcentuales entre estimadores se muestran en la variable dummy de Santiago.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-3,349	0,000	-3,390	0,000	-3,360	0,000
DummyStgo	-0,584	0,000	-0,459	0,000	-0,324	0,001
DummyCerrillos	-1,345	0,002	-1,411	0,000	-1,413	0,000
DummyCerroNavia	-0,312	0,082	-0,327	0,048	-0,326	0,051
DummyConchali	-0,823	0,000	-0,791	0,000	-0,735	0,000
DummyElBosque	-0,544	0,016	-0,568	0,002	-0,564	0,000
DummyEstacionCentral	-1,226	0,000	-1,219	0,000	-1,178	0,000
DummyIndependencia	-0,795	0,000	-0,835	0,000	-0,840	0,000
DummyLaCisterna	-0,327	0,009	-0,263	0,045	-0,166	0,267
DummyLaFlorida	-0,157	0,022	-0,237	0,006	-0,276	0,007
DummyLaGranja	-0,338	0,012	-0,211	0,128	-0,008	0,958
DummyLaReina	-0,285	0,013	-0,194	0,106	-0,119	0,387
DummyLasCondes	-1,062	0,000	-0,968	0,000	-0,839	0,000
DummyLoEspejo	-0,460	0,070	-0,452	0,028	-0,404	0,034
DummyLoPrado	-0,661	0,002	-0,506	0,005	-0,399	0,030
DummyMacul	-0,772	0,000	-0,698	0,000	-0,610	0,000
DummyMaipu	-0,377	0,000	-0,334	0,000	-0,284	0,006
DummyÑuñoa	-0,338	0,000	-0,273	0,005	-0,203	0,094
DummyPeñalolen	-0,808	0,000	-0,728	0,000	-0,658	0,000
DummyProvidencia	-1,179	0,000	-1,016	0,000	-0,903	0,000
DummyPudahuel	-0,419	0,008	-0,370	0,006	-0,295	0,025
DummyRecoleta	-0,937	0,000	-0,916	0,000	-0,859	0,000
DummyRenca	-0,382	0,023	-0,370	0,017	-0,286	0,065
DummySanJoaquin	-0,534	0,002	-0,506	0,001	-0,455	0,002
DummySanMiguel	-1,420	0,000	-1,258	0,000	-1,108	0,000
DummyVitacura	-0,345	0,002	-0,215	0,087	-0,069	0,662
HABITACIONAL	1,174E-05	0,019	1,429E-05	0,021	1,707E-05	0,032
OFICINA	0,000	0,019	0,000	0,007	0,000	0,024
SALUD	0,008	0,000	0,006	0,000	0,003	0,193
ln_fv	0,384	0,000	0,372	0,000	0,374	0,000
ln_fp	0,064	0,008	0,075	0,000	0,078	0,000
ln_fbTS	0,334	0,000	0,339	0,000	0,320	0,000
ln_transbordos	0,037	0,000	0,034	0,000	0,033	0,001

Tabla 6-31: Resumen de la modelación para la variable de accidentes no atropello, tipo 2.
Fuente: Elaboración propia.

El resumen de la Tabla 6-31 muestra los estimadores y su significancia para la variable explicada de accidentes no atropello, para los modelos tipo 2. Estos son cuando las variables incluidas son al menos significativas en una formulación (Normal, Poisson y Binomial Negativa). Los signos de los estimadores no varían entre formulaciones.

Las mayores diferencias porcentuales entre los estimadores se encuentran en las variables dummy de La Granja, dummy Vitacura y salud, en donde en esta última casi se triplica el estimador entre Binomial Negativo y Normal. La mayoría de las veces vuelven a ser menores los estimadores para el caso de la formulación Binomial Negativa.

Las variables no significativas son: salud, dummy La Cisterna y dummy Vitacura para los modelos Binomial Negativo; dummy La Graja y dummy La Reina para Poisson y Binomial Negativo.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-4,636	0,000	-5,063	0,000	-5,116	0,000
DummyLaFlorida	-0,559	0,001	-0,495	0,001	-0,475	0,002
DummyLasCondes	-0,641	0,000	-0,612	0,000	-0,602	0,000
DummyMacul	-0,688	0,031	-0,544	0,025	-0,479	0,041
DummyMaipu	-0,878	0,000	-0,745	0,000	-0,681	0,000
DummyPeñalolen	-0,473	0,041	-0,435	0,030	-0,419	0,035
DummySanJoaquin	-0,797	0,094	-0,625	0,026	-0,576	0,022
DummySanMiguel	-1,802	0,096	-1,730	0,000	-1,694	0,000
Vialidadkmsinautopistas	0,008	0,000	0,008	0,000	0,010	0,000
COMERCIO	0,000	0,001	0,000	0,006	0,000	0,007
INDUSTRIA	-0,003	0,001	-0,003	0,000	-0,003	0,000
CULTO	0,007	0,038	0,011	0,005	0,014	0,005
ln_fv	0,182	0,000	0,183	0,000	0,179	0,000
ln_fp	0,119	0,000	0,126	0,000	0,117	0,000
ln_fbTS	0,355	0,000	0,404	0,000	0,424	0,000
ln_transbordos	0,079	0,000	0,066	0,000	0,056	0,000

Tabla 6-32: Resumen de la modelación para la variable de accidentes atropello, tipo 1.
Fuente: Elaboración propia.

La variable explicada de accidentes atropello, para los modelos tipo 1 donde todas las variables son significativas en todas las formulaciones, se presentan en la Tabla 6-32. Todos los signos son los mismos entre las tres formulaciones.

La mayor diferencia porcentual entre los estimadores de las tres formulaciones se presenta en la variable comercio. La mayoría de las veces el modelo Binomial Negativo estima coeficientes en valor absoluto menores que el resto, y en el caso de la Tabla 6-32, el modelo Normal también lo hace para cinco parámetros.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-5,313	0,000	-5,349	0,000	-5,370	0,000
DummyCerrillos	-1,106	0,109	-1,286	0,006	-1,325	0,001
DummyLaFlorida	-0,586	0,000	-0,535	0,000	-0,514	0,001
DummyLasCondes	-0,312	0,060	-0,468	0,005	-0,515	0,004
DummyLoPrado	-0,437	0,036	-0,179	0,426	-0,085	0,730
DummyMacul	-0,689	0,026	-0,524	0,027	-0,474	0,039
DummyMaipu	-0,936	0,000	-0,787	0,000	-0,715	0,000
DummyPeñalolen	-0,392	0,090	-0,383	0,052	-0,380	0,050
DummyRecoleta	-0,604	0,004	-0,342	0,071	-0,247	0,208
DummySanJoaquin	-0,806	0,076	-0,622	0,024	-0,579	0,020
DummySanMiguel	-1,842	0,081	-1,743	0,000	-1,691	0,000
Vialidadkmsinautopistas	0,008	0,000	0,009	0,000	0,011	0,000
COMERCIO	0,000	0,000	0,000	0,002	0,000	0,004
INDUSTRIA	-0,003	0,003	-0,002	0,005	-0,002	0,003
CULTO	0,011	0,001	0,010	0,014	0,012	0,016
SALUD	0,007	0,000	0,005	0,045	0,003	0,405
SumadeAvalúoFiscaldelospredios	-8,431E-13	,000	-2,208E-13	,083	-1,422E-13	,236
In_fv	0,254	0,000	0,230	0,000	0,225	0,000
In_fp	0,196	0,000	0,158	0,000	0,143	0,000
In_fm	-0,066	0,000	-0,052	0,000	-0,045	0,004
In_fbTS	0,310	0,000	0,391	0,000	0,411	0,000
In_transbordos	0,080	0,000	0,060	0,000	0,050	0,001

Tabla 6-33: Resumen de la modelación para la variable de accidentes atropello, tipo 2.
Fuente: Elaboración propia.

La Tabla 6-33 muestra el resumen de los modelos para la variable de accidentes del tipo atropello, para los modelos en donde las variables son incluidas si son significativas en, al menos, una formulación (tipo 2).

Nuevamente los signos son todos iguales entre formulaciones. La mayoría de las veces el estimador del caso Binomial Negativo es menor que para las otras 2 formulaciones. En cuanto a las diferencias porcentuales entre los estimadores, se dan para el caso de la suma de avalúo fiscal de predios y para la variable de salud, en donde el estimador por Binomial Negativo es casi 2,5 veces mayor que el estimador del modelo Normal.

Las variables que no son significativas son: dummy de Cerrillos para el modelo Normal; dummy Lo Prado para el modelo Poisson y Binomial Negativo; dummy Recoleta, salud y suma de avalúo fiscal de predios para la formulación Binomial Negativa.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-3,277	0,000	-3,736	0,000	-3,808	0,000
DummyEIBosque	0,689	0,000	0,697	0,000	0,702	0,000
DummyLaGranja	0,983	0,000	1,057	0,000	1,244	0,000
DummyLaPintana	0,914	0,000	0,859	0,000	0,804	0,000
DummyLaReina	0,416	0,002	0,433	0,002	0,391	0,012
DummyÑuñoa	0,739	0,000	0,729	0,000	0,713	0,000
DummyPAC	0,657	0,000	0,643	0,000	0,651	0,000
DummyQtaNormal	1,029	0,000	1,032	0,000	1,041	0,000
DummySanMiguel	-1,404	0,058	-1,431	0,000	-1,438	0,000
DummyVitacura	0,466	0,000	0,512	0,000	0,567	0,001
DummyPuenteAlto	0,665	0,000	0,594	0,000	0,445	0,000
In_fv	0,304	0,000	0,322	0,000	0,342	0,000
In_fp	0,059	0,017	0,070	0,003	0,068	0,001
In_fbTS	0,297	0,000	0,329	0,000	0,311	0,000
In_transbordos	0,041	0,001	0,034	0,003	0,039	0,001

Tabla 6-34: Resumen de la modelación para la variable de lesionados, tipo 1.
Fuente: Elaboración propia.

La modelación para la cantidad de lesionados, según los modelos tipo 1 donde todas las variables incluidas son significativas, se presentan en la Tabla 6-34. Se puede notar que los signos de todos los estimadores no varían entre formulaciones.

La mayor diferencia porcentual se presenta en la variable dummy de la comuna Puente Alto y, a diferencia de los modelos anteriores para accidentes totales, no atropello y atropello, los menores valores absolutos de los coeficientes estimados los calcula la formulación Normal.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-2,876	0,000	-3,367	0,000	-3,618	0,000
DummyCerrillos	-1,502	0,172	-1,530	0,000	-1,546	0,000
DummyEIBosque	0,606	0,000	0,607	0,000	0,615	0,000
DummyLaGranja	0,935	0,000	1,004	0,000	1,184	0,000
DummyLaPintana	0,822	0,000	0,766	0,000	0,712	0,000
DummyLaReina	0,368	0,005	0,365	0,007	0,288	0,059
DummyMaipu	-0,206	0,147	-0,210	0,083	-0,240	0,043
DummyÑuñoa	0,691	0,000	0,659	0,000	0,629	0,000
DummyPAC	0,610	0,000	0,597	0,000	0,608	0,000
DummyQtaNormal	0,971	0,000	0,978	0,000	0,999	0,000
DummyRecoleta	-0,432	0,094	-0,420	0,024	-0,401	0,011
DummySanMiguel	-1,442	0,049	-1,474	0,000	-1,483	0,000
DummySanRamon	-0,466	0,292	-0,529	0,074	-0,540	0,015
DummyVitacura	0,448	0,000	0,484	0,000	0,509	0,003
DummyPuenteAlto	0,590	0,000	0,497	0,000	0,335	0,004
HABITACIONAL	7,022E-06	,232	1,349E-05	,060	2,590E-05	,005
In_fv	0,275	0,000	0,295	0,000	0,321	0,000
In_fp	0,052	0,038	0,059	0,012	0,053	0,014
In_fbTS	0,291	0,000	0,329	0,000	0,333	0,000
In_transbordos	0,044	0,000	0,035	0,002	0,034	0,002

Tabla 6-35: Resumen de la modelación para la variable de lesionados, tipo 2.
Fuente: Elaboración propia.

En la Tabla 6-35 se muestra el resumen de los estimadores y su significancia para las tres formulaciones de los modelos de la cantidad de lesionados como variable explicada. Corresponden a los modelos de tipo 2 donde las variables incluidas son al menos significativas en una formulación.

Los signos de los estimadores, al igual que todos los modelos presentados anteriormente, no varían entre Normal, Poisson y Binomial Negativo. Las mayores diferencias que los coeficientes poseen están en la variable habitacional, donde es casi 4 veces menor el estimador del modelo Normal versus el modelo Binomial Negativo.

Al igual que el caso de la misma variable explicada de cantidad de lesionados de la Tabla 6-34, en general los menores valores absolutos de los estimadores se calculan mediante la formulación Normal.

Las variables no significativas de este tipo de modelos, entre las tres formulaciones, fueron: dummy Cerrillos, dummy Maipú, dummy San Ramón y habitacional, todas en el modelo Normal. Para el resto de los parámetros en las distintas formulaciones, fueron todas significativas por sobre un 90% de confianza.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-5,726	0,000	-5,776	0,000	-5,838	0,000
DummyLaPintana	1,126	0,000	1,022	0,001	0,980	0,007
DummyRenca	0,883	0,001	0,756	0,047	0,721	0,075
DummySanJoaquin	0,884	0,000	0,768	0,027	0,738	0,047
DummyPuenteAlto	0,699	0,000	0,672	0,007	0,665	0,014
HABITACIONAL	3,836E-05	,004	4,049E-05	,037	4,050E-05	,057
In_fbTS	0,619	0,000	0,628	0,000	0,638	0,000

Tabla 6-36: Resumen de la modelación para la variable de fallecidos, tipo 1.
Fuente: Elaboración propia.

Para la variable de cantidad de fallecidos por zona, en la Tabla 6-36 se muestra el resumen de las tres formulaciones para los modelos tipo 1, donde todas las variables son significativas.

Los signos nuevamente no varían entre formulaciones.

La mayor diferencia porcentual la presenta la variable dummy Renca entre Binomial Negativo y Normal. Además, en cuatro de los siete parámetros el modelo binomial estima los menores valores absolutos de los coeficientes. En los otros tres casos lo hace el modelo normal.

Parámetro	NORMAL		POISSON		BINOMIAL NEGATIVO	
	Estimador	Sig.	Estimador	Sig.	Estimador	Sig.
(Intersección)	-6,156	0,000	-6,223	0,000	-6,293	0,000
DummyLaPintana	1,362	0,000	1,194	0,000	1,145	0,003
DummyQtaNormal	0,634	0,051	0,575	0,134	0,562	0,156
DummyRenca	1,016	0,000	0,829	0,030	0,784	0,052
DummySanJoaquin	0,947	0,000	0,828	0,017	0,798	0,031
DummyPuenteAlto	0,861	0,000	0,761	0,003	0,742	0,006
HABITACIONAL	3,488E-05	,011	3,831E-05	,053	3,857E-05	,073
In_fv	0,139	0,078	0,111	0,183	0,109	0,196
In_fbTS	0,485	0,000	0,537	0,000	0,550	0,000

Tabla 6-37: Resumen de la modelación para la variable de fallecidos, tipo 2.
Fuente: Elaboración propia.

La Tabla 6-37 muestra los coeficientes estimados para la variable de cantidad de fallecidos en modelos del tipo 2, donde las variables incluidas son significativas en al menos una de las tres formulaciones. Los signos no varían entre los coeficientes.

Los estimadores son en seis de los 9 parámetros menores en valor absoluto en los modelos Binomial Negativo. Los otros tres casos son menores para la formulación Normal.

Las dos variables no significativas son la dummy de Quinta Normal para el caso de los modelos Poisson, y la variable del logaritmo neperiano del flujo vehicular (automóviles) en la formulación Binomial Negativa.

Otra manera de observar el ajuste y los análisis de residuos es calculando el error medio absoluto, o mean absolute error (MAE) en inglés. Esta medida representa el valor absoluto de la diferencia entre lo observado y lo predicho, dividido por la cantidad de observaciones que se tiene. Su fórmula se puede observar en la Ecuación 17. Se puede interpretar como la media del error de la modelación, sin considerar si se subestima o se sobrestima la cantidad de accidentes.

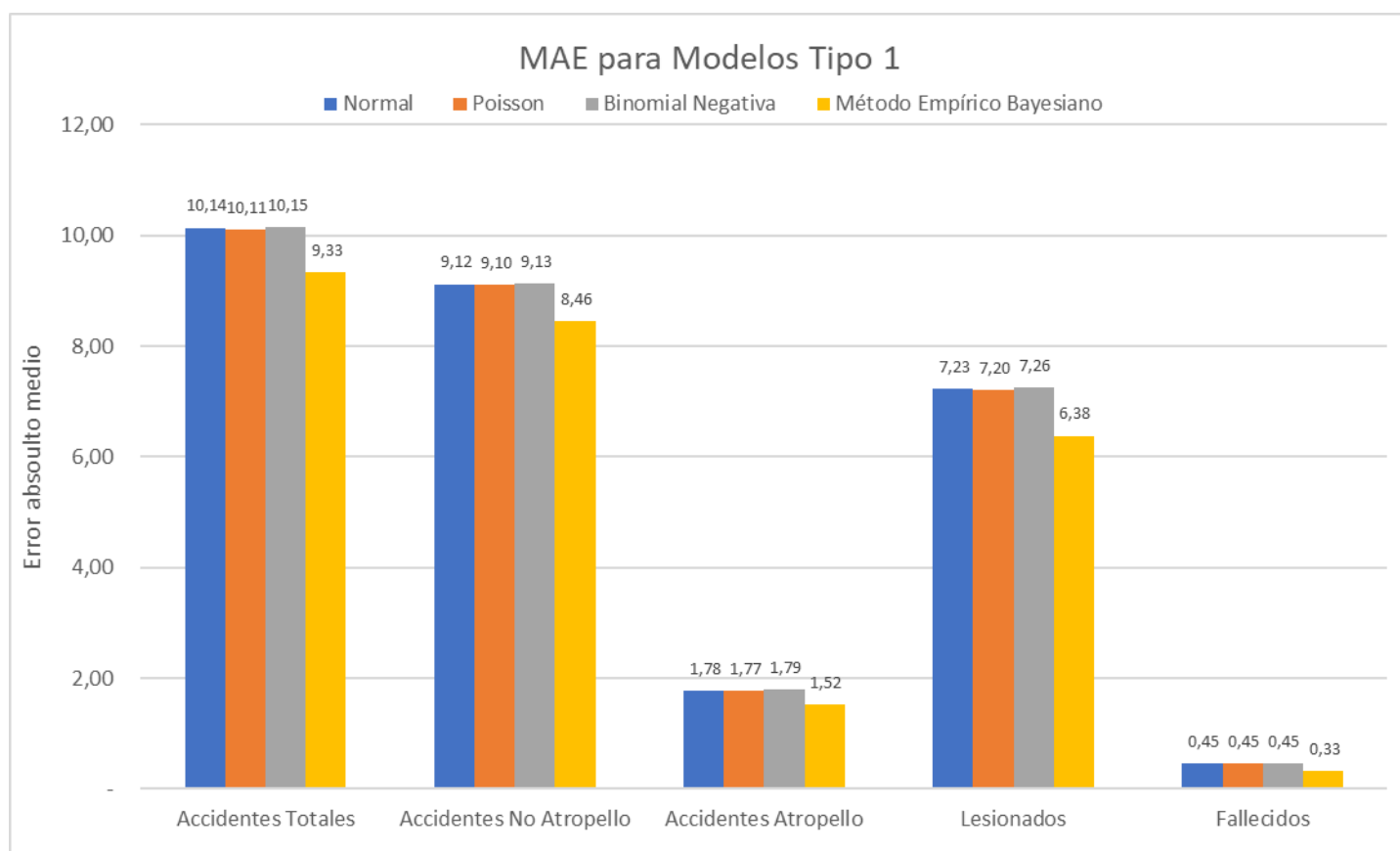
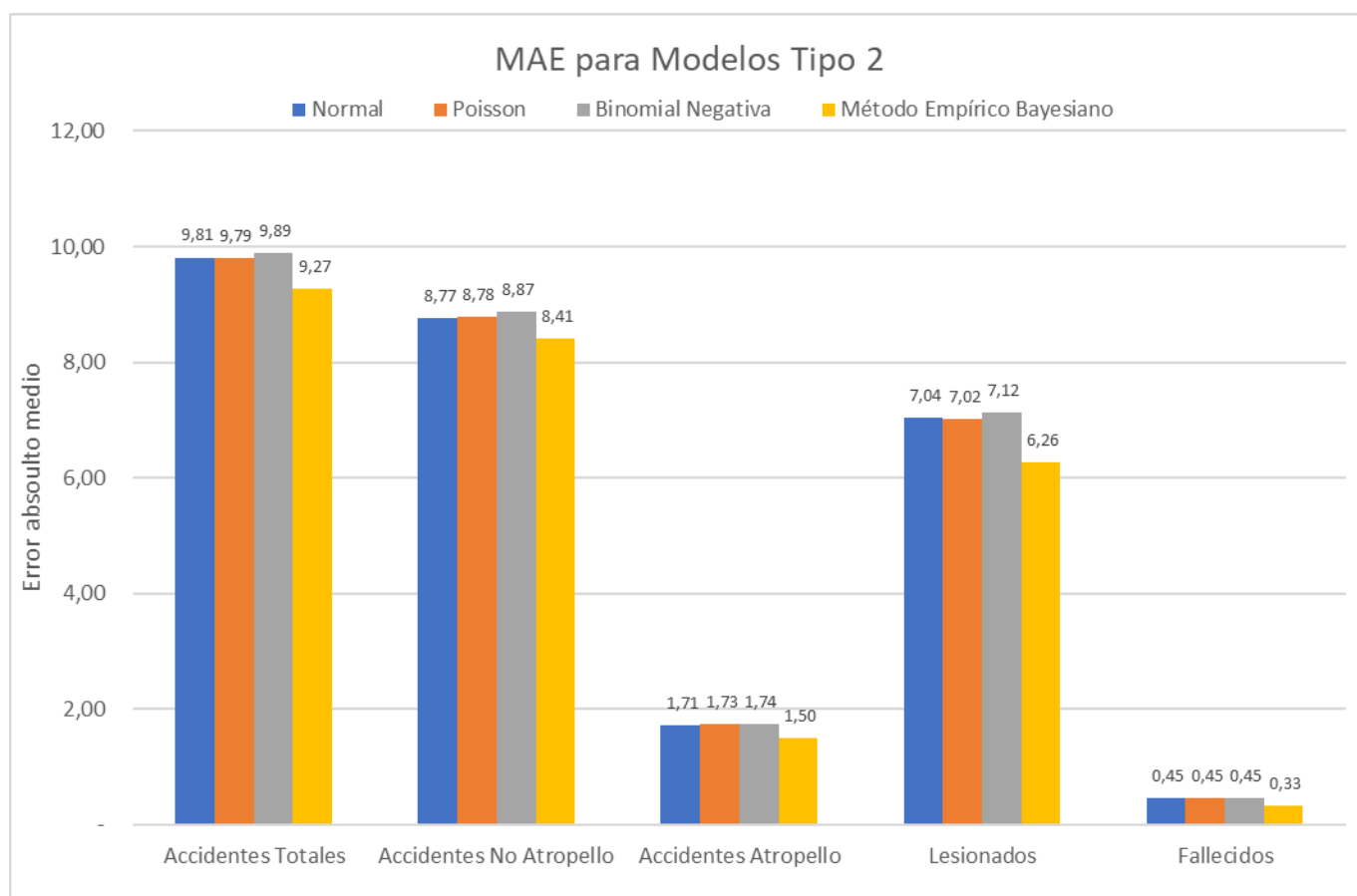


Gráfico 7: Estadístico MAE para la modelación 2013, tipo 1.

Fuente: Elaboración propia.



*Gráfico 8: Estadístico MAE para la modelación 2013, tipo 2.
Fuente: Elaboración propia.*

Se puede observar del Gráfico 7 y el Gráfico 8 que el menor valor del error absoluto medio, sin considerar el método empírico bayesiano, lo presenta, en general, la formulación Poisson. Esto para los modelos de accidentes totales, accidentes no atropello tipo 1, accidentes atropello tipo 1 y cantidad de lesionados. En todo el resto de los casos es la modelación normal la que se ajusta de mejor forma a la realidad.

Cabe destacar que en todos los casos la estimación mediante el Método Empírico Bayesiano posee menor error absoluto medio.

7 Estimación de tasas de accidentabilidad para el año 2015

Para poder estimar la tasa de accidentalidad con los modelos, primero se modifican las variables incluidas en la modelación. Posteriormente se realiza la estimación para las tres formulaciones (Normal, Poisson y Binomial Negativa). Luego se calcula la predicción mediante el método empírico bayesiano en el caso de la formulación binomial negativa. Estos resultados se comparan con lo registrado para el año 2015. Cabe destacar que el método empírico Bayesiano considera un registro histórico, el cual es el calculado como promedio de las variables explicadas entre los años 2011 y 2014.

7.1 Variaciones en las variables independientes

Para poder realizar una comparación entre lo que predicen los modelos y la realidad, se generan las actualizaciones de las variables importantes en la modelación y luego se calculan las tasas predichas según los parámetros estimados, reportados anteriormente.

Las variables de flujo vehicular, de motos, peatonal y la cantidad de transbordos fueron actualizadas según la variación que se constata de las encuestas de movilidad para la ciudad de Santiago del año 2001 y 2012, reportadas en la EOD 2012 (SECTRA, 2014) y resumidas en la Tabla 7-1.

Modo de Transporte	2001		2012		Diferencia	
	Viajes (miles)	%	Viajes (miles)	%	Diferencia (miles)	%
Auto	3.275.964	21,0%	4.586.676	26,1%	1.310.712	40,0%
Bip!	4.690.452	30,1%	4.164.335	23,7%	-526.117	-11,2%
Bip! - Otros Público	384.062	0,8%	419.994	1,4%	35.932	98,1%
Bip! - Otros Privado	189.099	0,4%	311.876	0,8%	122.777	123,7%
Taxi Colectivo	121.054	2,5%	239.780	2,4%	118.726	9,4%
Taxi Básico	65.094	1,2%	145.605	1,8%	80.511	64,9%
Otros	566.469	3,6%	923.977	5,3%	357.507	63,1%
Caminata	5.964.457	38,3%	6.075.651	34,6%	111.194	1,9%
Bicicleta	328.979	2,1%	676.007	3,9%	347.028	105,5%
Total	15.585.630	100,0%	17.543.901	100,0%	1.958.271	12,6%

Tabla 7-1: Evolución de la participación modal 2001 – 2012.
Fuente: (SECTRA, 2014).

Con lo anterior se genera una regresión lineal entre los años y la cantidad de viajes, para luego obtener una tasa de aumento entre el año 2012 y el 2015. Las tasas obtenidas, que representan también el factor de ampliación por el cual se va a aumentar o disminuir los flujos, se presentan en la Tabla 7-2.

Modo	Factor de ampliación de flujos al año 2015
Auto	1,0779
Bici	1,1400
Peatonal	1,0050
Moto	1,1055
Transbordos	1,0321

Tabla 7-2: Factor de ampliación de los flujos al año 2015 por modo.
Fuente: Elaboración propia.

Por otro lado, dado que se cuenta con los programas de operación de Transantiago mediante lo publicado por el DTPM, se procede a calcular nuevamente el flujo de buses utilizando la misma metodología explicada en el apartado 5.4.

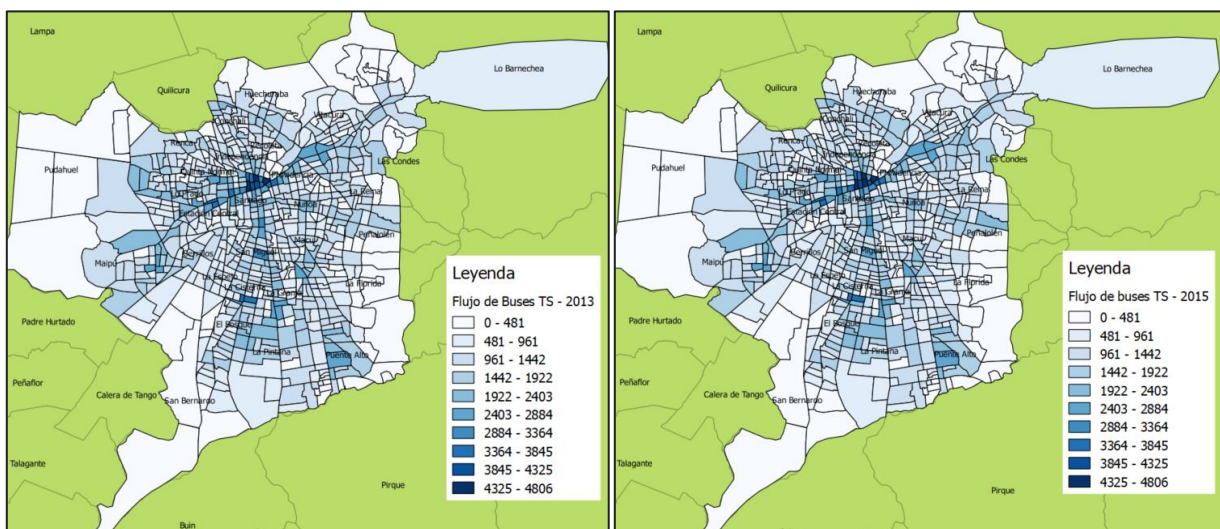


Figura 14: Flujo de buses por zona para el año 2013 (izquierda) y para el año 2015 (derecha).
Fuente: Elaboración propia.

Se puede notar que existen diferencias mínimas entre el flujo de buses en los 2 años mostrados en la Figura 14. La media del flujo para el año 2013 es de 1.147,4 buses por día promedio de la semana y su desviación estándar es de 711,3. Para el año 2015 su media es ligeramente menor (1.137,5) y su desviación estándar ligeramente mayor (715,0).

Por otro lado, las variables explicativas se obtienen de lo antes mostrado en el apartado 5.1 donde se cuenta con información de la base de datos SIEC-2 trabajada por CONASET, en particular para el año 2015.

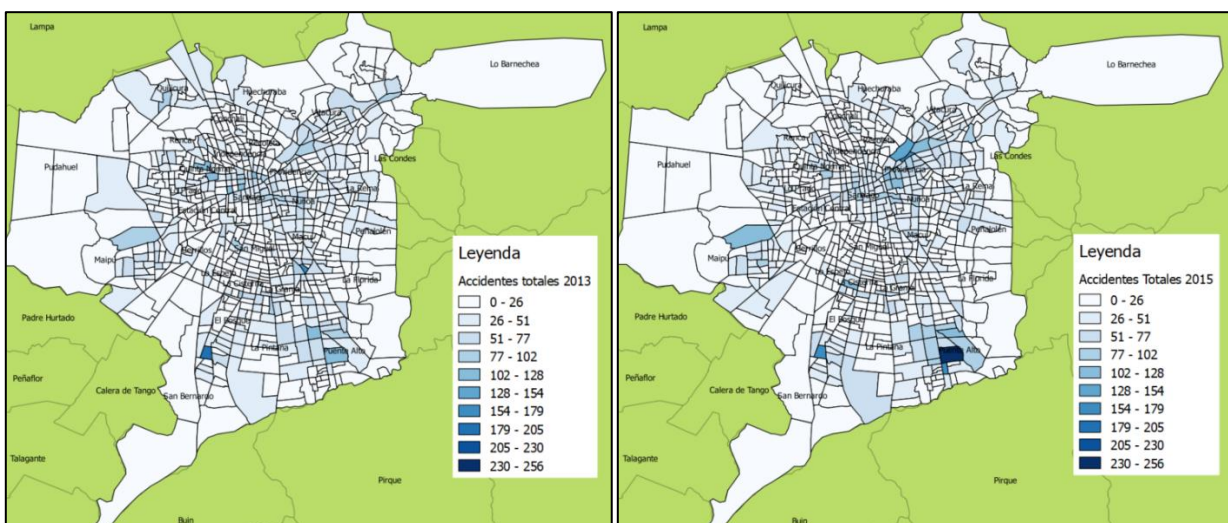


Figura 15: Cantidad de accidentes totales en el año 2013 (izquierda) y 2015 (derecha).
Fuente: Elaboración propia.

Se puede notar de la Figura 15 una diferencia en la cantidad de accidentes reportados entre los años 2013 y 2015. Particularmente en sectores de la comuna de Puente Alto y la comuna de Providencia.

Por otro lado, para ajustar la cantidad de predios, se consiguió la base de datos de la cantidad de predios para los 2 semestres del 2015, pero no se obtuvieron archivos georreferenciados para poder cruzarlo con la información de la zonificación. Por esta razón es que se decide analizar la cantidad de predios por uso de suelo y comparar los datos del 2013 al respecto del 2015 para luego amplificar por un factor que corresponde a la variación por uso de suelo entre estos años.

Uso de suelo	Factor de amplificación
COMERCIO	1,0120
DEPORTE Y RECREACION	1,0210
EDUCACION CULTURA	1,0155
HOTEL MOTEL	1,0574
HABITACIONAL	1,0435
INDUSTRIA	0,9938
BIENES COMUNES	1,0402
BODEGA Y ALMACENAJE	1,1193
MINERIA	1,0000
OFICINA	1,0929
ADMINISTRACION PUBLICA Y DEFENSA	1,0163
CULTO	1,0198
SALUD	1,0121
TRANSPORTE Y TELECOMUNICACIONES	1,0227
OTROS NO CONSIDERADOS	1,0480
SITIO ERIAZO	1,0386
ESTACIONAMIENTO	1,1303

Tabla 7-3: Factor de amplificación por uso de suelo para el año 2015.
Fuente: Elaboración propia.

Se puede observar de la Tabla 7-3 que el único factor que disminuye la cantidad de predios es el de industria. El mayor factor es el de estacionamiento, pero esta variable no es utilizada en el análisis. Cabe notar que las variables que sí se utilizan son: comercio, habitacional, industria, oficina, culto y salud. De los factores correspondientes a estos usos de suelo, el mayor es el de oficina.

Otra variable incluida en la modelación es la suma de avalúo fiscal de los predios. Para extender esta se realizó un proceso análogo al anterior. El factor de amplificación utilizado fue 1,1352 calculado como la suma de avalúo fiscal de todos los predios del 2015, dividido por su correspondiente valor del 2013.

Dentro de las variables utilizadas en los modelos se encuentra la de vialidad por zona sin incluir las autopistas. Para la predicción de accidentes del año 2015 se considera que la vialidad no se altera, ya que no se tienen las herramientas suficientes ni los medios para poder analizar en el tiempo dicha variable. Además, es esperable que, en dos años a nivel zonal, las variaciones en vialidad sean pequeñas.

7.2 Comparación de lo predicho con lo observado

Para mostrar los resultados de la predicción, se obtienen los valores residuales. Estos corresponden a la diferencia entre lo real y lo predicho. Con estos datos se construyen histogramas para observar qué tan alejados están del valor 0 que corresponde a la predicción certera. Además, se calcula el estadístico de la media del error absoluto al estimar con los modelos antes presentados.

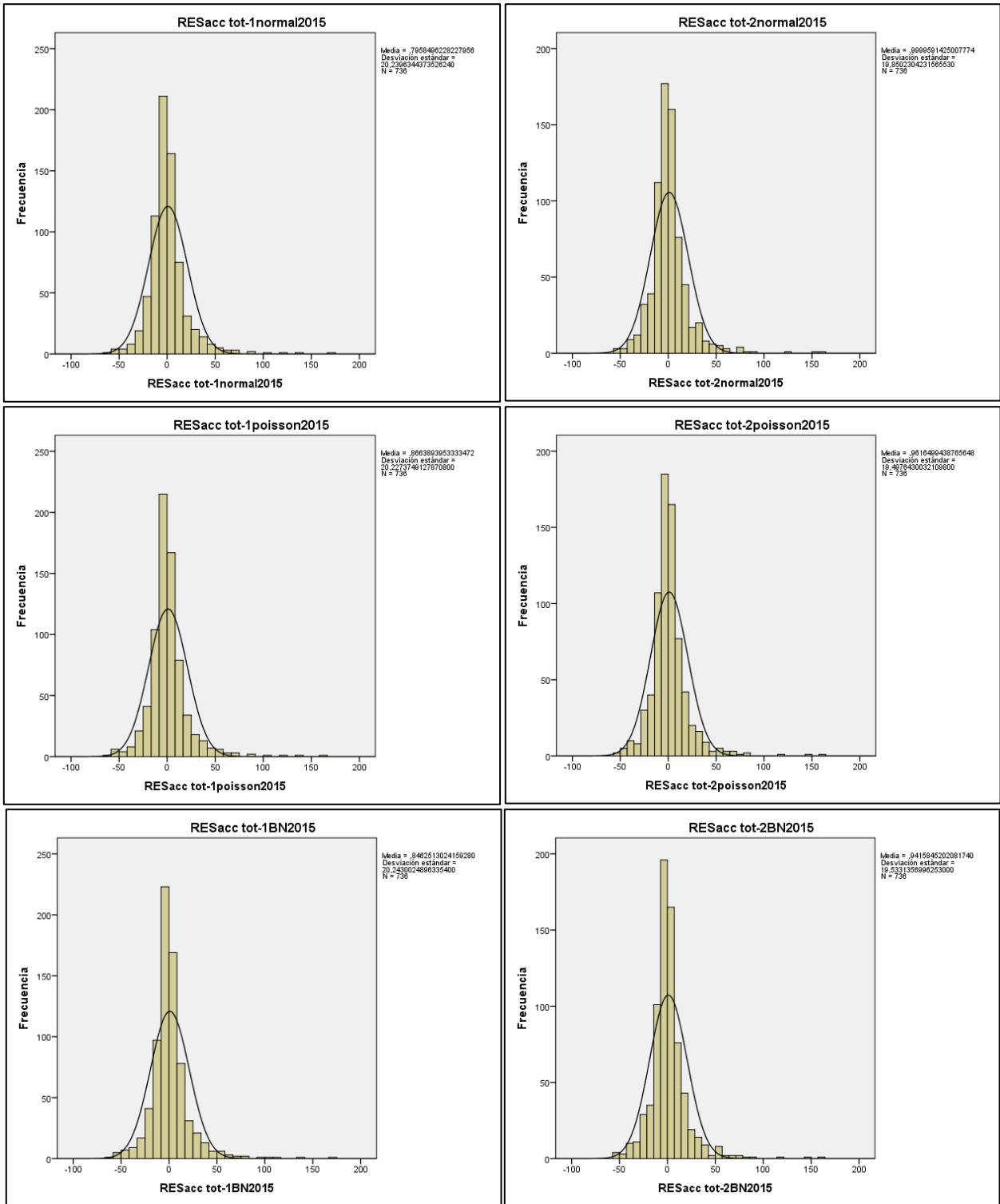


Gráfico 9: Histogramas de los residuos para la modelación de tasas de accidentes totales del año 2015.
Fuente: Elaboración propia.

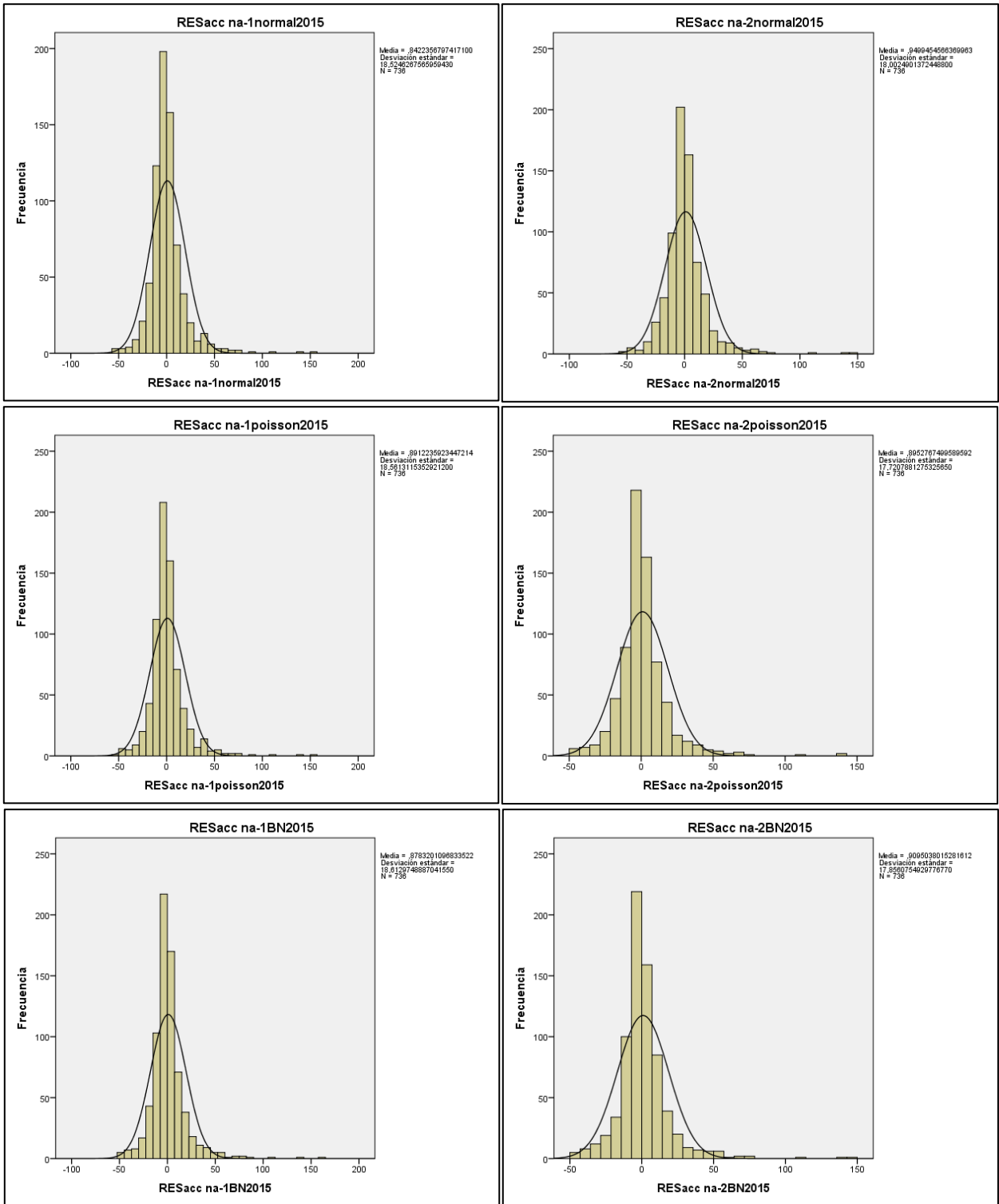


Gráfico 10: Histogramas de los residuos para la modelación de tasas de accidentes tipo no atropello del año 2015. Fuente: Elaboración propia.

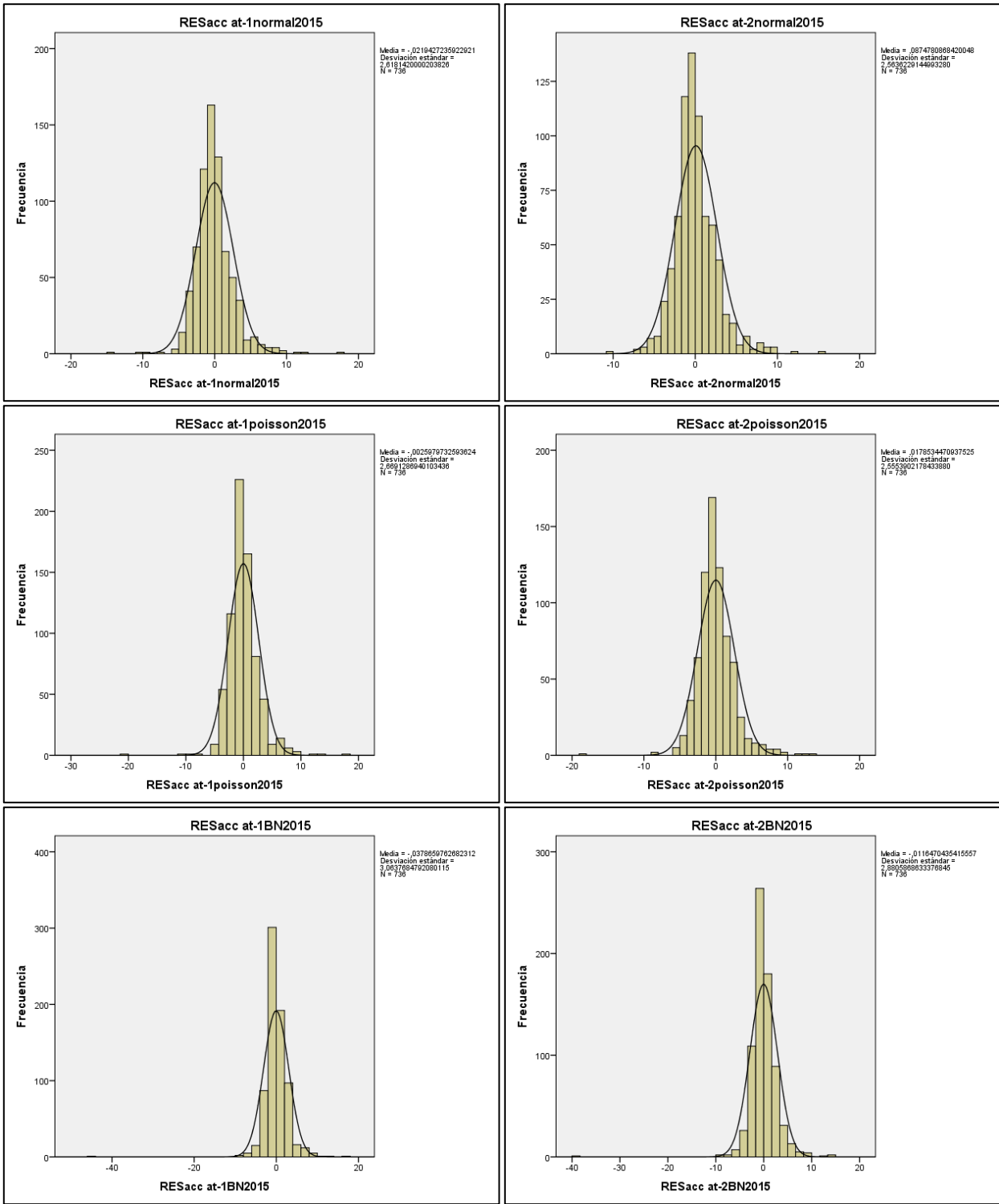


Gráfico 11: Histogramas de los residuos para la modelación de tasas de accidentes tipo atropello del año 2015. Fuente: Elaboración propia.

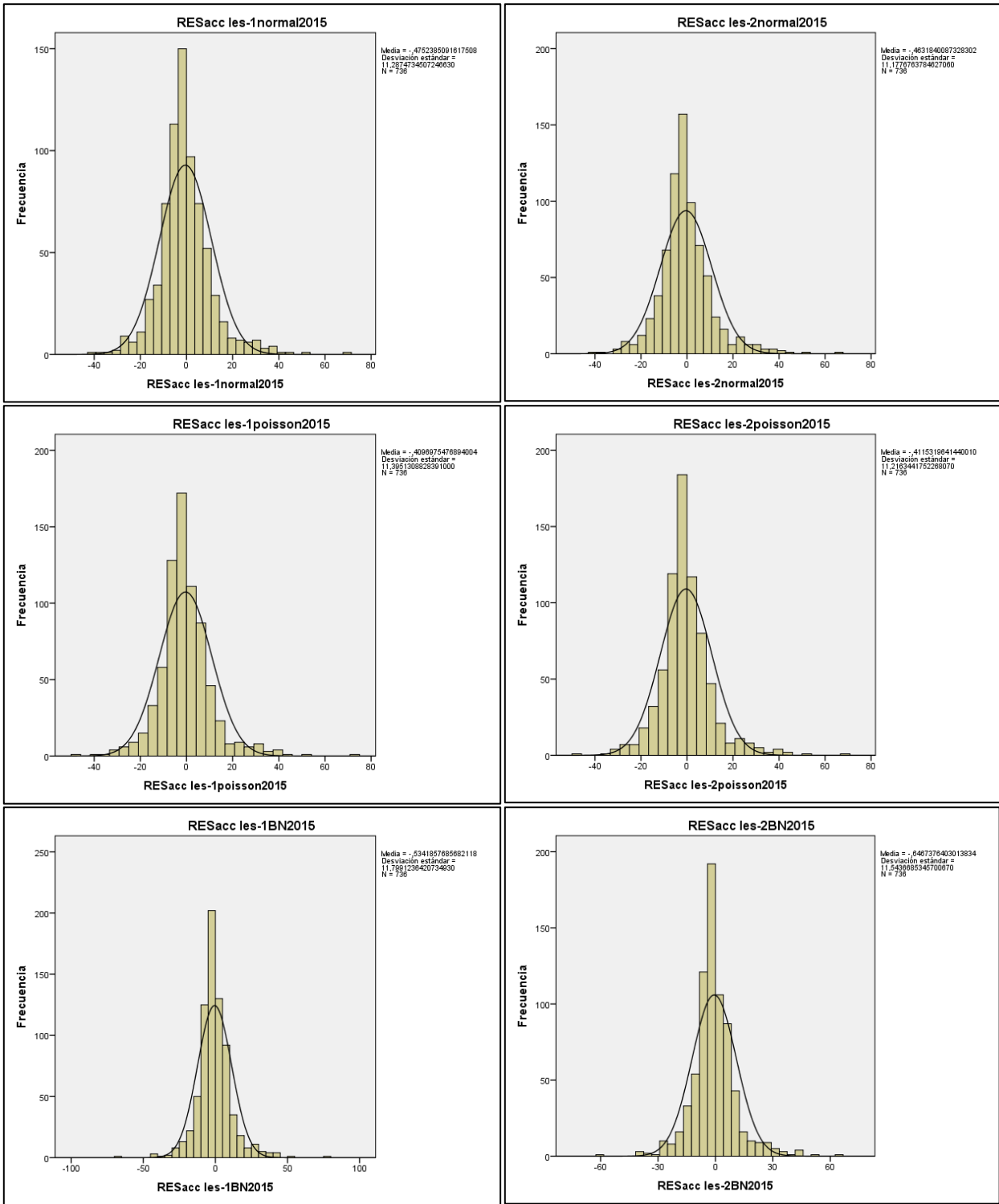


Gráfico 12: Histogramas de los residuos para la modelación de lesionados del año 2015.
Fuente: Elaboración propia.

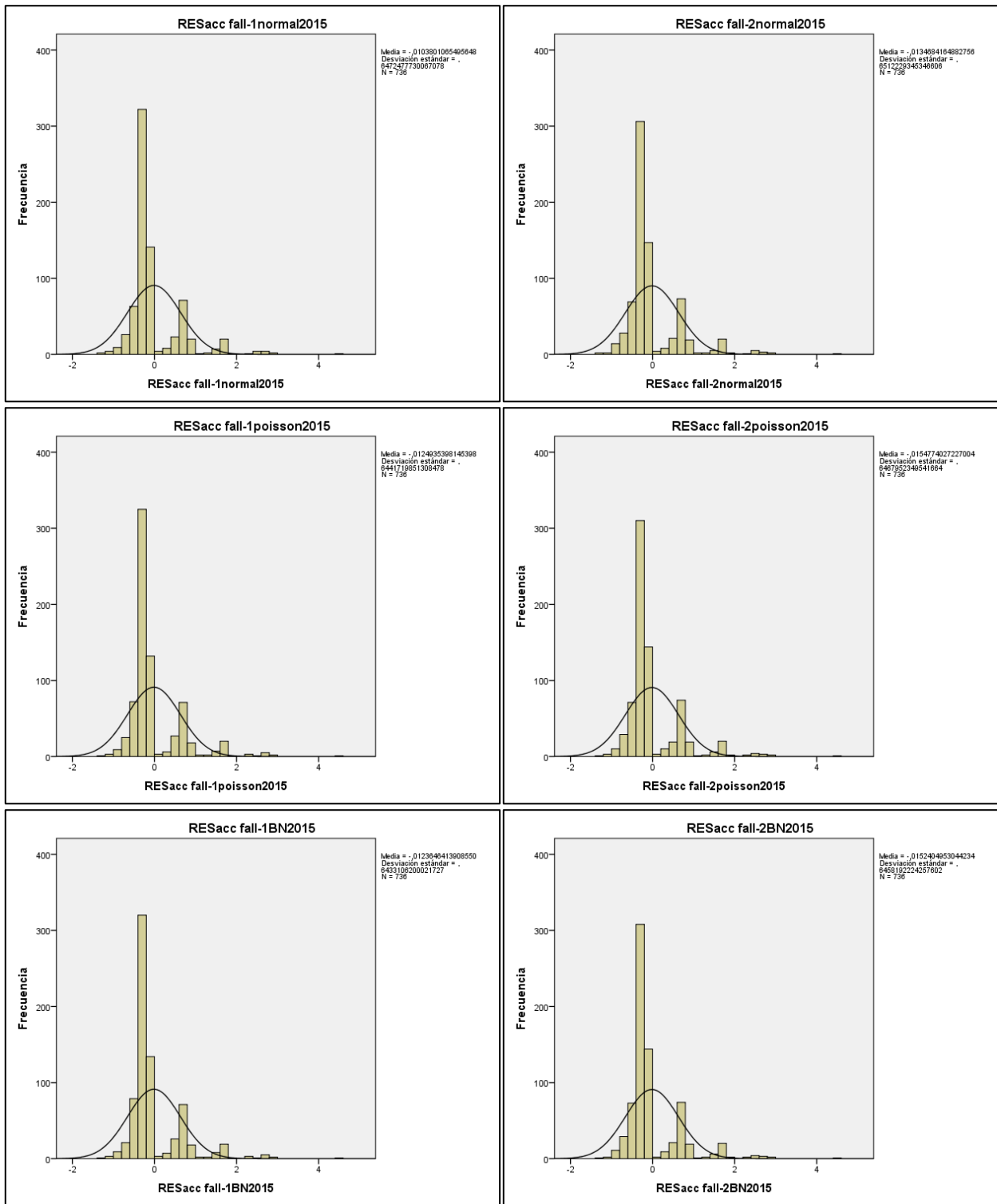


Gráfico 13: Histogramas de los residuos para la modelación de fallecidos del año 2015.
Fuente: Elaboración propia.

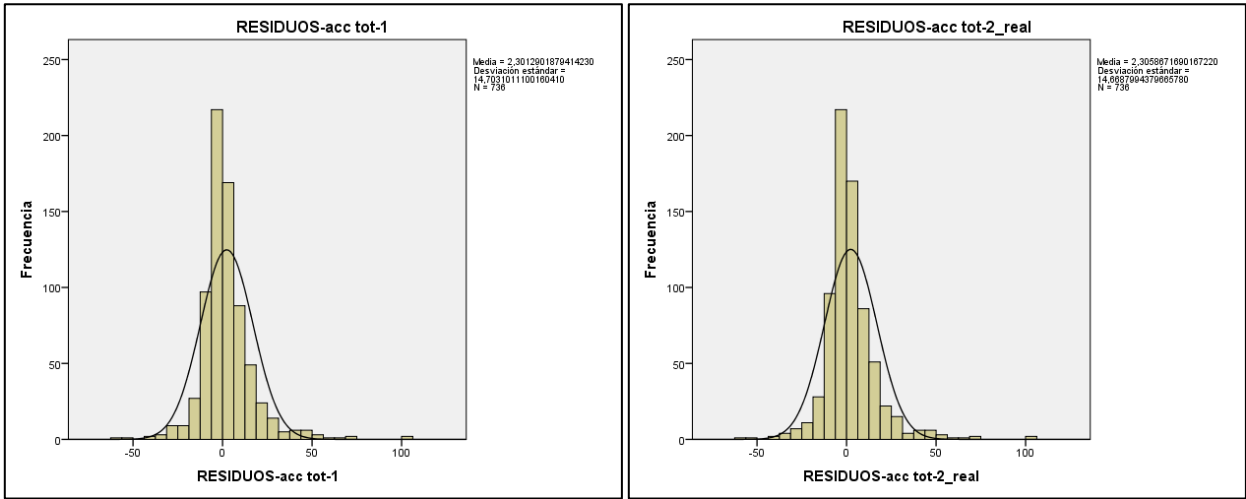


Gráfico 14: Histogramas de los residuos del método empírico bayesiano de los accidentes totales.
 Fuente: Elaboración propia.

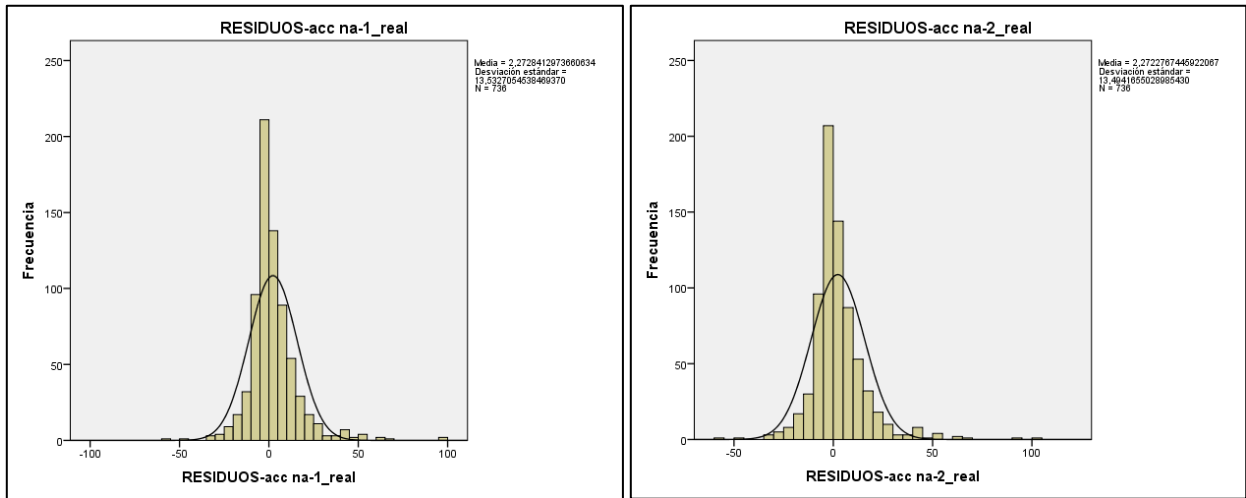


Gráfico 15: Histogramas de los residuos del método empírico bayesiano de los accidentes no atropello.
 Fuente: Elaboración propia.

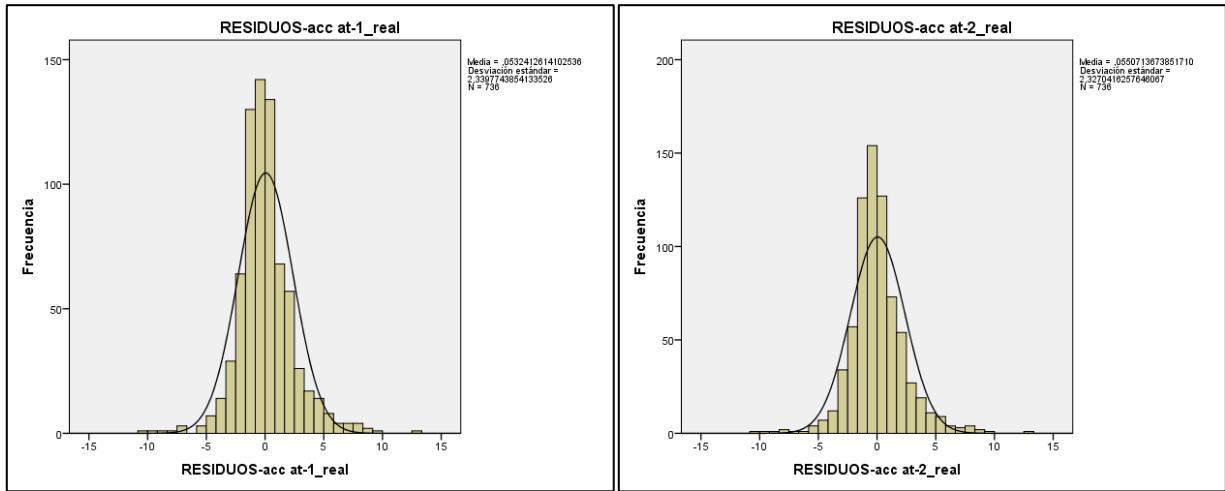


Gráfico 16: Histogramas de los residuos del método empírico bayesiano de los accidentes atropello.
Fuente: Elaboración propia.

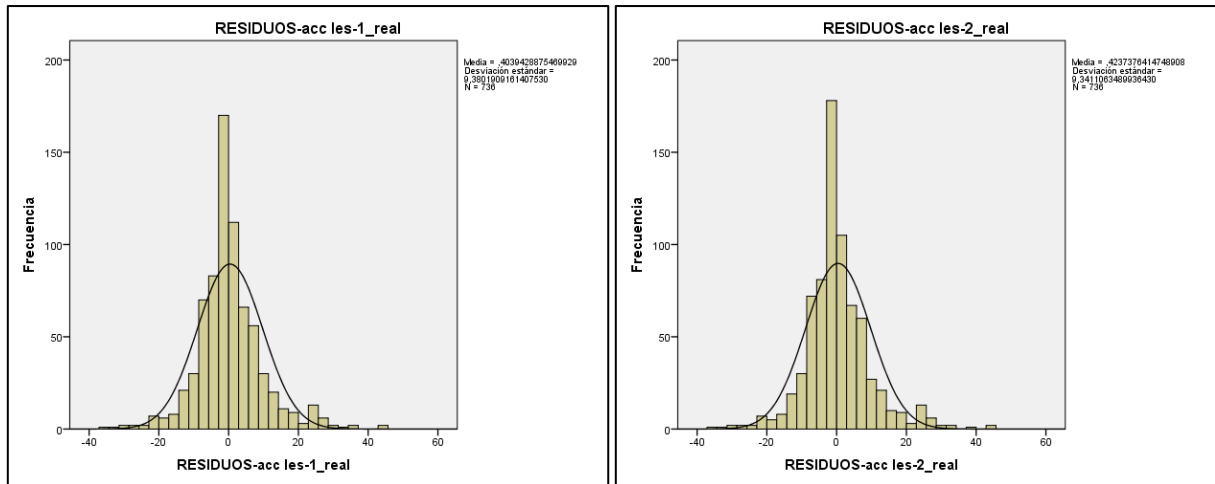


Gráfico 17: Histogramas de los residuos del método empírico bayesiano de los lesionados.
Fuente: Elaboración propia.

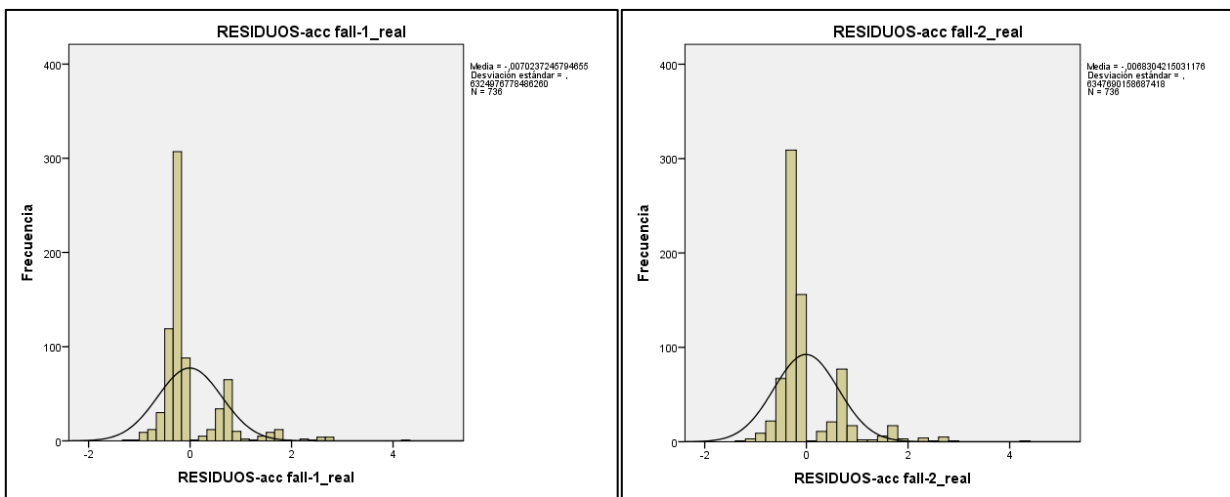


Gráfico 18: Histogramas de los residuos del método empírico bayesiano de los fallecidos.
Fuente: Elaboración propia.

De los modelos para las tasas de accidentes totales en el Gráfico 9 se puede observar que el modelo que tiene una menor media es el Normal-1, sin considerar el método empírico bayesiano. Las desviaciones estándar son relativamente parecidas, al igual que la media, variando desde 19,50 hasta 20,24 accidentes por zona por año (acc/zona-año).

En los modelos de accidentes tipo no atropello del Gráfico 10 el que tiene mejor valor de la media residual, de las 3 formulaciones principales, es el Normal-1. La desviación varía entre 17,72 y 18,61 acc/zona-año, ligeramente menor que el caso anterior.

Para los accidentes tipo atropello en el Gráfico 11, el que menor valor de la media tiene es Poisson-1 con valor -0.003 acc/zona-año, sustancialmente cercana a cero, lo cual es ideal. Sus desviaciones estándar van desde 2,56 y 3,06 acc/zona-año.

Los residuos en los modelos de cantidad de lesionados en el Gráfico 12 presentan valores siempre negativos, lo que sugiere que se subestima la cantidad personas heridas, sin tener en cuenta el método empírico bayesiano. El modelo de menor valor de media residual es Poisson-1 con -0.41 acc/zona-año. Las desviaciones estándar están entre 11,18 y 11,80 acc/zona-año.

Para los modelos de la cantidad de fallecidos por zona en el Gráfico 13 se puede observar que los residuos no siguen una tendencia parecida a la curva gaussiana normal. La menor media es de -0,010 acc/zona-año y las desviaciones varían entre 0,6433 y 0,6512 acc/zona-año.

Observando ahora los histogramas de los residuos para el método empírico bayesiano, de los modelos de accidentes totales en el Gráfico 14 se puede notar que la media es superior a cero, lo que indica que se sobreestima la frecuencia de siniestros. La desviación estándar es 14,70 y 14,67 acc/zona-año y las medias tienen casi el mismo valor: 2,30 y 2,31 acc/zona-año.

En el caso de los residuos de los accidentes tipo no atropello para el método empírico bayesiano en el Gráfico 15 sucede una situación similar a la anterior. La media es de 2,27 acc/zona-año y la desviación estándar es 13,49 y 13,53 acc/zona-año.

En los modelos de accidentes tipo atropello con dicho método en el Gráfico 16, sucede algo distinto y más alentador. La media es cercana a cero con valores de 0,05 y 0,06 acc/zona-año. La desviación estándar es de 2,34 y 2,33 acc/zona-año.

Para la cantidad de lesionados en el Gráfico 17 la media es de 0,40 y 0,42 acc/zona-año. El error estándar es de 9,38 y 9,34 acc/zona-año.

En el caso de los fallecidos del Gráfico 18, al igual de lo que sucede con los residuos de la modelación sin aplicar el método empírico bayesiano, la media es muy cercana a cero, tomando el valor de -0,007 acc/zona-año. La desviación típica en este caso es de 0,63 acc/zona-año.

Otra forma de observar el nivel de ajuste de la predicción es comparar lo predicho con lo que se observa para el año 2015 mediante el estadístico del error absoluto medio MAE de la Ecuación 17. Los valores para los distintos modelos se presentan en el Gráfico 19 y el Gráfico 20.

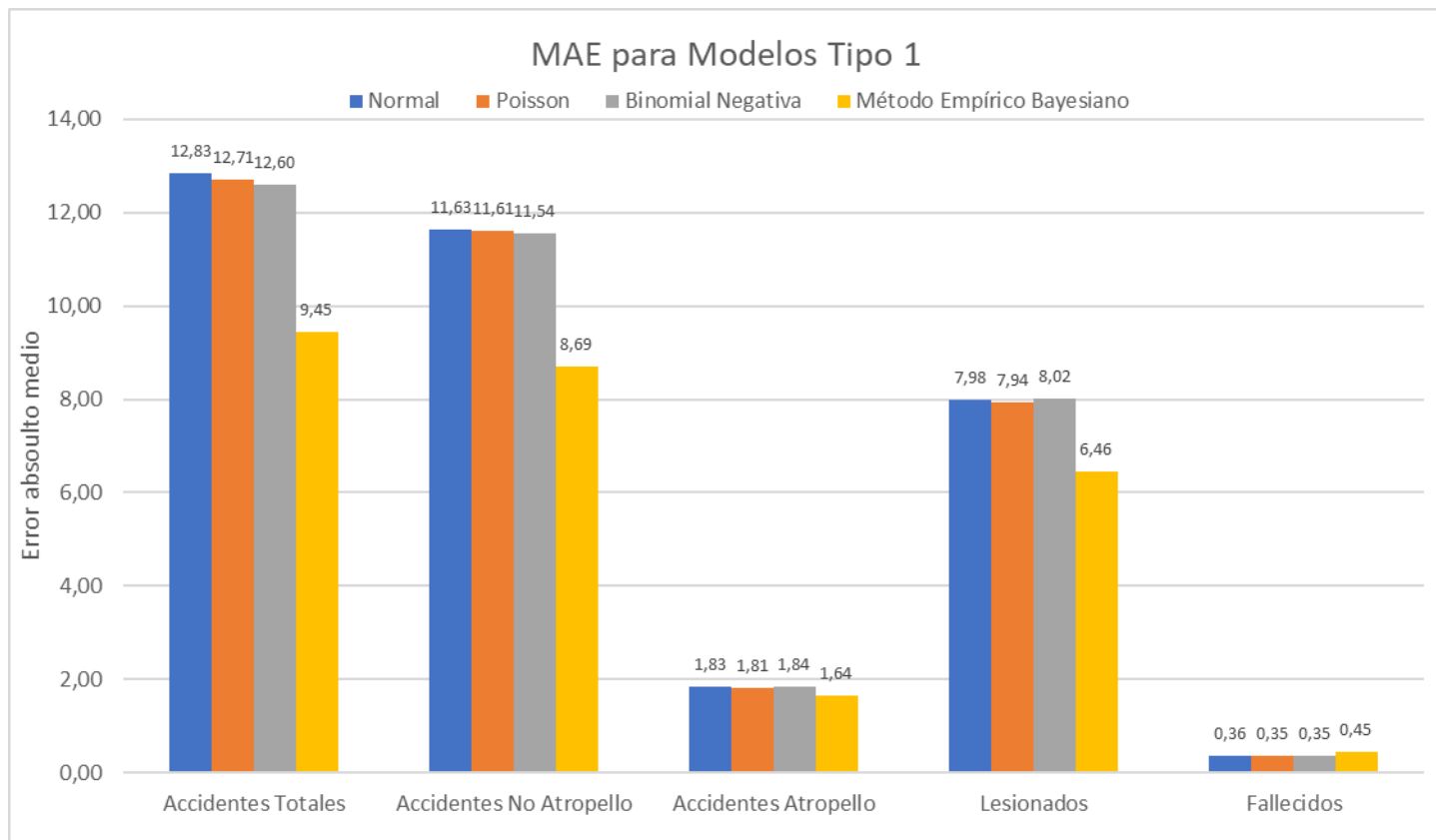


Gráfico 19: Estadístico MAE para la predicción 2015 de los modelos tipo 1.
Fuente: Elaboración propia.

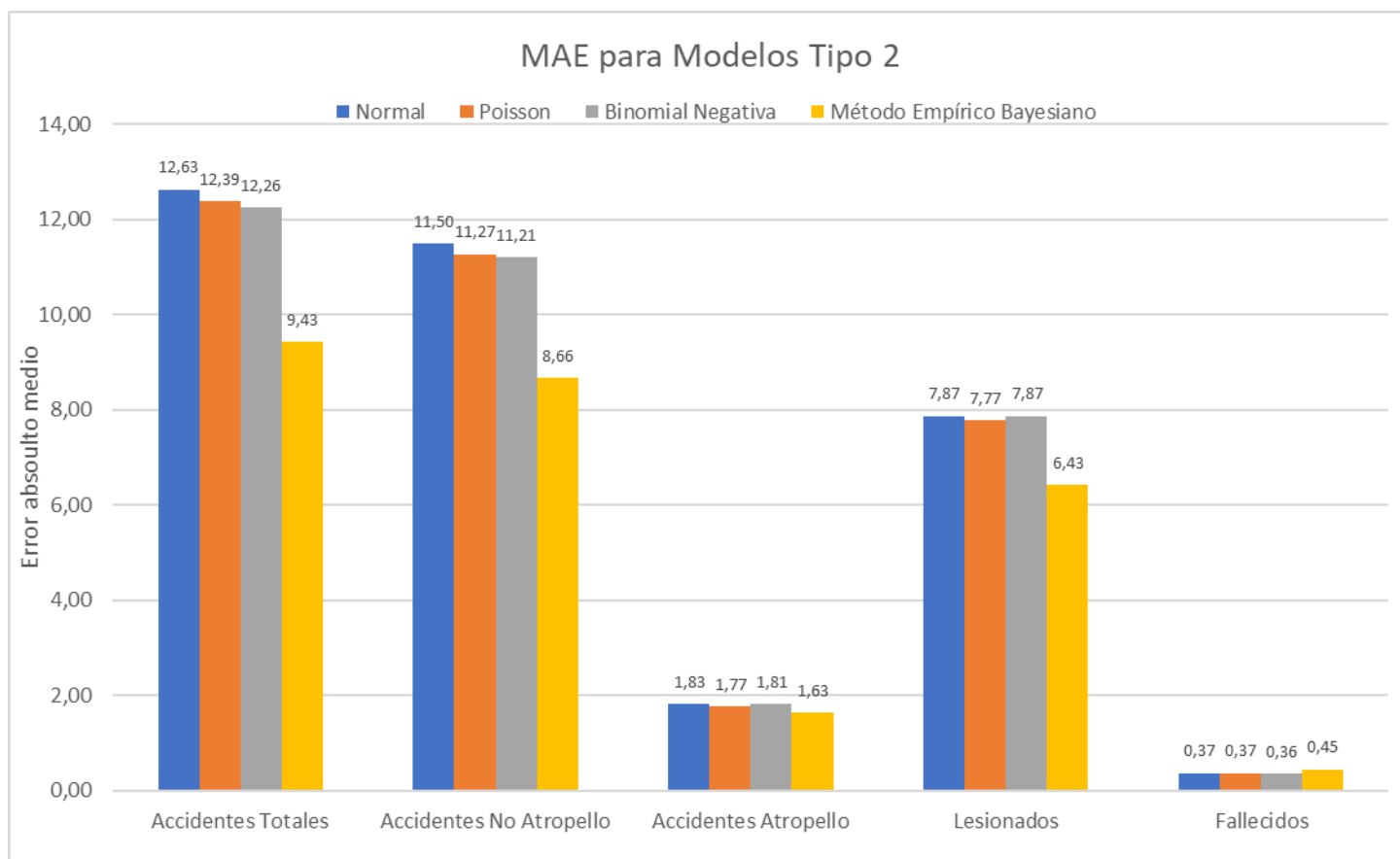


Gráfico 20: Estadístico MAE para la predicción 2015 de los modelos tipo 2.
Fuente: Elaboración propia.

En el Gráfico 19 y el Gráfico 20 se puede observar el estadístico MAE para todas las formulaciones de todos los modelos con distintas variables explicadas. Además, se presenta el estadístico del error medio absoluto para el método empírico bayesiano.

Se puede notar entonces que el error medio de la predicción de accidentes en los modelos de accidentes totales es ligeramente menor para el caso de los modelos binomial negativo, y sustancialmente menor para el método empírico bayesiano. El mismo patrón se repite para los modelos de accidentes tipo no atropello.

Distinto es el caso de los accidentes atropello y para lesionados, donde si bien la formulación Poisson tiene un mejor estadístico MAE, el método empírico bayesiano es menor y por lo tanto estima de mejor forma los accidentes de este tipo.

Para el caso de los modelos de la cantidad de fallecidos, el método empírico bayesiano posee un estadístico del error medio absoluto mayor que el caso de las formulaciones Normal, Poisson y Binomial Negativa. Para los modelos tipo 1 y los modelos tipo 2, la mejor estimación las realiza el modelo Binomial Negativo.

Cabe notar también que los menores valores de la media de errores absolutos se presentan para los modelos tipo 2, donde se incluyen más variables que explican los fenómenos, aunque algunas de ellas no sean significativas.

8 Conclusiones

8.1 Descripción del trabajo realizado

En esta memoria de título, se aborda la modelación de los accidentes de tránsito en Santiago con tres formulaciones matemáticas, buscando comparar los resultados y estimar un modelo que se ajuste lo mejor posible a la realidad. Estas formulaciones corresponden a modelos lineales generalizados Normal, Poisson y Binomial Negativo. Este último, considera el fenómeno de la regresión a la media y la sobredispersión de los datos. Además, se utiliza el método empírico bayesiano para comprobar su ajuste en modalidad predictiva.

Se utilizan variables que explican el fenómeno según lo observado en la literatura y según la disponibilidad de los datos existentes para la ciudad de Santiago. Las fuentes de información corresponden a la base de datos de accidentes de Carabineros de Chile SIEC-2, además de la base de datos de bienes raíces del SII, Encuesta de movilidad Origen Destino 2012 del Ministerio de Transporte y Telecomunicaciones, Programas de operación del sistema de transporte público Transantiago, modelo de simulación basada en agentes a gran escala para el flujo vehicular y aforos obtenidos por DICTUC, entre otras.

El trabajo de procesamiento de los datos consistió, principalmente, en cruzar la información georreferenciada. Se generaron agregaciones espaciales de las fuentes de datos a un nivel de zona definido en la Encuesta de movilidad Origen Destino 2012. Estos procesos incluyeron análisis de rutas mínimas, corrección de información faltante o errada, cálculos de tasas y estadísticos que se desprenden de las fuentes de datos, estadística espacial e histórica, además del proceso de recopilación de información relevante incluso considerando bases de datos que no fueron utilizadas.

Junto con una revisión de la literatura científica, se plantea una serie de predictores que explican el fenómeno de los siniestros viales. Estos predictores tienen relación con la cantidad de predios y su destino de uso de suelo por cada zona de modelación donde destacan los usos habitacionales, oficina, industria, culto, salud y comercio. Además, se utiliza la suma del avalúo fiscal de los predios, variables de flujo en distintos modos (Auto, bicicleta, caminata, bus del transporte público y motocicleta), transbordos, vialidad y variables tipo dummy de pertenencia de una zona a una comuna. Sobre esto último, en las 736 zonas de modelación se consideran 33 ayuntamientos incluyendo Puente Alto y San Bernardo, además de los de la provincia de Santiago, excluyendo Quilicura.

Las variables explicadas que se modelaron fueron: accidentes totales, accidentes tipo no atropello, accidentes tipo atropello, cantidad de lesionados y cantidad de fallecidos. Se crean 2 tipos de modelos para estas variables: tipo 1 donde todas las variables explicativas son significativas y otro, tipo 2, donde las variables incluidas son estadísticamente significativas en al menos una de las tres formulaciones.

Se observan los estadísticos Desviación, chi cuadrado de Pearson, logaritmo de la verosimilitud, AIC y el error absoluto medio para comparar las formulaciones y los tipos de modelos en modalidad predictiva para el año 2015 y ajustando los parámetros a lo observado en el año 2013.

8.2 Resultados

Los resultados principales se presentan a continuación.

Sobre la modelación y la comparación entre las tres formulaciones de este trabajo, en los modelos basados en una distribución Normal se destaca el hecho de que, en general, no presentan un mejor ajuste en sus estadísticos para la modalidad predictiva, en comparación con las otras dos formulaciones. Son capaces de ajustarse mejor a la realidad en base a lo observado por el estadístico MAE, que contradice lo observado por los estadísticos AIC o por el logaritmo de la verosimilitud. Esto puede indicar que, si bien el valor predicho se acerca mucho más a lo observado, no se cumplen los supuestos de la modelación y, por lo tanto, los modelos no son correctamente utilizados.

Sobre los modelos Poisson, el sesgo que implica no considerar la sobredispersión afecta los cálculos de los estimadores y, por lo tanto, puede estar sesgando los efectos reales de las variables en la modelación. Esto se puede notar en que, en modalidad predictiva, los modelos Poisson en general no entregan mejores ajustes que la binomial negativa, excepto en el caso de los accidentes del tipo atropello y cantidad de lesionados. La mejor estimación de residuos, en general, la entrega el modelo Poisson al modelar el año 2013, sin embargo, en modalidad predictiva son los modelos binomiales negativa los que se ajustan de mejor forma al año 2015.

Sobre estos últimos modelos (BN), se puede notar que, al utilizar el método empírico bayesiano, los residuos mejoran su capacidad predictiva y entregan una mayor cantidad de aciertos, excepto para los modelos de cantidad de fallecidos. Se puede destacar que disminuye entre un 0,15 y un 3,15 el estadístico MAE. Sería interesante analizar este mismo efecto para distintos años y así poder afirmar esta conclusión para otros periodos.

Una de las consideraciones de los modelos lineales generalizados es la independencia entre las observaciones. Esto puede no ser del todo cierto en el caso de los accidentes en una medida pequeña cuando, por ejemplo, consideramos un accidente en el que produjo congestión o efectos en el tráfico inesperados que los usuarios no alcanzan a prever y por lo tanto produce más siniestralidad. Esto es esperable que suceda en pocos casos, pero afectarán la independencia de las observaciones. En este contexto, se puede establecer que existen conductores con mayor aversión al riesgo que otros y, por lo tanto, algunos están más expuestos a participar de un siniestro vial, siendo este un fenómeno imperceptible por los modelos aquí presentados.

La principal conclusión estadística de este trabajo es relativa al ajuste de los modelos Binomial Negativo por sobre los otros dos, ya que estos presentan mejores estadísticos

de bondad de ajuste, predicen mejor la realidad al observar sus estadísticos, explican y consideran el fenómeno de regreso a la media y la sobredispersión natural que se tienen en los datos de seguridad vial. La aplicación de los modelos tipo 2 junto con el método empírico bayesiano muestra el mejor ajuste posible alcanzado en el desarrollo de esta memoria de título y sugieren su utilización en la planificación de las medidas a favor de la seguridad vial.

Sobre los signos de los estimadores, la variable habitacional junto con los flujos de automóviles, peatones, buses y transbordos, tienen siempre estimadores positivos, indicando que a mayor cantidad de predios o de flujo, mayor será la tasa de accidentes. Esta conclusión es esperable dado que el flujo indica el nivel de exposición, y la variable habitacional sugiere que, a mayor cantidad de hogares, mayor será la exposición al tránsito.

Se puede notar también que, para los modelos de accidentes totales, accidentes no atropello y atropellos, todas las variables dummy tienen estimadores de signo negativo, excepto para la comuna de San Ramón en la modelación de los accidentes totales, tipo 2. Esto quiere decir que las comunas que se incluyen en la modelación van a registrar menor cantidad de accidentes que para el resto de las zonas. Esto no se condice con lo esperado, por ejemplo, para el caso de la comuna de Santiago, en donde su variable dummy se considera para la modelación de accidentes totales y para los accidentes no atropello. Podría explicarse esto debido a que estas variables pueden estar recogiendo efectos exógenos que no han sido percibidos por las demás variables. Además, se destaca que esta variable no es significativa en el análisis de los atropellos, lesionados y fallecidos, que podría indicar que la cantidad alta de accidentes registrados en esta comuna sean debido a las variables de modelación incluidas, como los flujos peatonales, vehiculares y de buses además de las variables de uso de suelo, o bien el fenómeno no se afecta por la pertenencia a esta comuna.

Por otro lado, las variables dummy que resultan significativas en los modelos de lesionados y fallecidos poseen estimadores positivos. Es decir que al pertenecer una zona a las comunas consideradas se va a aumentar el riesgo de siniestralidad. Es importante notar entonces que la comuna de La Pintana es significativa en la modelación de la cantidad de lesionados y de fallecidos, así como la variable para Quinta Normal. Los registros indican que en la comuna de La Pintana la cantidad de lesionados es de 302 y 13 fallecidos, para el año 2013, en 368 accidentes totales. En el mismo año, para Quinta Normal, se reconocen 562 lesionados y 9 fallecidos en 700 accidentes totales. Estas cifras no son las más negativas en cuanto a la siniestralidad de la ciudad de Santiago, pero sí representan un valor por sobre el promedio de accidentes totales (544) excepto en La Pintana, de lesionados (330) y de fallecidos (7,2) por comuna.

Respecto de las variables y su correlación, es importante destacar el hecho de que varias de ellas poseen un estadístico de correlación de Pearson mayor a 0,5 que se incluyen en la modelación y que al momento de elegir qué variables incluir y cuáles no, jugó un rol importante. Por ejemplo, el uso de suelo oficina fue probado para ver si la variable comercio era significativa al incluirla. En todos los casos la variable de oficina sí fue significativa y comercio no, excepto en el caso de la modelación para los accidentes del

tipo atropello. Esto se podría explicar debido a que el comercio genera mayor cantidad de viajes de peatones de lo que lo hace la oficina, en diversos rangos horarios, y por lo tanto la exposición de peatones es mayor, implicando un aumento en la siniestralidad. La variable oficina tiene estimadores negativos y el comercio positivo.

Es importante también notar que, la variable de flujo peatonal y el flujo de bicicletas poseen alta correlación, por lo que se prueba la significancia de ambas, siendo siempre prevalente la de peatones. Esto permite afirmar la conclusión de que el flujo de peatones sí es significativo para la modelación de la tasa de accidentes.

También hay variables de uso de suelo que están correlacionadas entre sí. Particularmente relevante resulta que los predios con destino de uso de suelo habitacional estén correlacionados con educación y cultura, bienes comunes y con culto. Esto indica que esta variable debiese ser significativa al momento de modelar, y efectivamente resulta en aquello, siendo significativa en todos los modelos, excepto en los atropellos.

Respecto de las correlaciones del nivel de flujo, la variable de buses del Transantiago resulta poseer correlación con las variables explicadas, que confluye en una alta significancia al momento de modelar las tasas de accidentes, como se puede observar en todos los modelos donde fue incluida. Es más, es la única variable que resultó significativa en los modelos de las variables explicadas de accidentes totales, accidentes no atropello, accidentes atropello, cantidad de lesionados y cantidad de fallecidos, para todas las formulaciones.

Dentro de los análisis presentados, se destaca el de la influencia de una persona en la siniestralidad para los modos de automóvil, bus del Transantiago y caminata. La conclusión es que, en casi todas las zonas, un viajero en modo auto tendrá más probabilidad de disminuir la seguridad vial para los modelos de accidentes totales y de accidentes no atropello, luego una persona en bus y finalmente el menor riesgo se presenta para una persona caminando. Por otro lado, el aumento porcentual que se registran para los modelos de accidentes atropello y lesionados indican que, en casi todas las zonas, el efecto marginal de una persona en bus será el de mayor efecto en el aumento de la siniestralidad. A su vez, una persona en auto representa mayor aumento porcentual de las tasas que una persona más viajando como peatón. Esto se nota como efecto marginal, por lo que es relevante que la tasa de ocupación promedio de un bus sea tan alta, versus la de un auto o un peatón. Además, se debe tener en consideración que se utilizaron tasas de ocupación constante, lo que sugiere indagar en esta información y calcular con mayor certeza estos efectos marginales.

Se condice entonces con la intuición el resultado obtenido para accidentes totales y para accidentes no atropello, siendo esta que el vehículo privado representa una mayor amenaza para la seguridad vial que el transporte público, y que es necesario tener en cuenta las relaciones del peatón con el tránsito para poder disminuir las consecuencias y las tasas de accidentabilidad. Por el contrario, un pasajero más en buses del transporte público representará el mayor riesgo en los modelos de accidentes atropello y lesionados, seguido por el riesgo de una persona más en automóvil y finalmente el peatón que

aumenta en la menor medida las tasas de siniestros. Esto se puede explicar debido a que en estos tipos de modelos la variable de flujo de buses puede estar percibiendo fenómenos que el resto de las variables no, y, por lo tanto, estar sobrestimando el efecto en relación con la implicancia de un automovilista más. Como se menciona anteriormente, las tasas de ocupación utilizadas fueron constantes para toda la ciudad, por lo que esto puede implicar una fuente de error importante. Finalmente, otra posible razón del porqué el efecto marginal de una persona más en bus es mayor que para un automovilista y un peatón, es que la participación de los buses en los accidentes tipo atropello y la cantidad de lesionados casi se duplica al respecto de los accidentes totales y los accidentes tipo no atropello, es decir, casi se duplica el riesgo en más de 8 veces los kilómetros recorridos por auto en vez de un bus. Cabe destacar también que, a modo de ejemplo, para el año 2012, fallecieron 4 personas cada 100 accidentes tipo atropello donde se vieron involucrados automóviles, y fallecieron 12 personas cada 100 atropellos donde se vieron involucrados buses.

8.3 Desafíos Futuros

En Chile, el Ministerio de Transporte se encarga de realizar estudios de movilidad urbana en diversas ciudades del país. Este trabajo se centró en la ciudad de Santiago por la disponibilidad de información y por el interés de considerar la información de la EOD 2012 que no había sido utilizada anteriormente para modelar este tipo de fenómenos.

En ese contexto es entonces desafiante considerar otras ciudades donde se dispone de la información utilizada en estos modelos, como lo son las encuestas de movilidad para estimar flujos peatonales, flujos de buses y flujos vehiculares, junto con la información del SII de bienes raíces e información de accidentes de tránsito donde se debiese presentar otro desafío importante: la georreferenciación de los accidentes de tránsito.

Es de esperar que el tratamiento de los datos realizado por CONASET sea revelador en este ámbito, por lo que queda abierta la posibilidad de ajustar estos modelos a las realidades regionales particulares, en donde también sería interesante tratar de incluir otras variables como el clima, la composición vehicular, composición del diseño vial, planimetrías, regulación, entre otras.

Es también notable el análisis en años posteriores cuando se tenga información nueva sobre los datos recopilados en este trabajo para poder observar si existieron mejoras significativas en materia de seguridad vial, si el fenómeno permanece estancado o en crecimiento. Podría ser importante también considerar los factores de contexto económico y la realidad país, tal cual lo indica la experiencia internacional, donde dicta que el factor temporal se correlaciona con tendencias y medidas ambientales exógenas de los factores, por ejemplo, el producto interno bruto, desempleo e indicadores macroeconómicos.

Por último, las desagregaciones propuestas y la modelación planteada están basadas en la información disponible y su calidad de reporte. Esto debiese mejorar con el tiempo

puesto que las voluntades de las autoridades y los organismos relativos a la seguridad vial han mostrado avances en esta materia.

Es de esperar que, junto con la mayor riqueza de información, la seguridad vial en el país tome la importancia que merece a todo nivel. La opinión pública presione la agenda política del país, ojalá más temprano que tarde, y los profesionales envueltos en los rubros de transporte puedan cambiar el paradigma hacia una visión de protección e impedimento de la escalada en los fallecimientos que el país lamenta hace muchos años.

9 Bibliografía

- AASHTO. (2010). Highway Safety Manual. Washington, DC. Retrieved from <http://www.highwaysafetymanual.org/Pages/default.aspx>
- Box, G., & Cox, D. (1981). *An Analysis of Transformations Revisited, Rebutted*: WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER.
- Camus, L. (2017). *Desarrollo y aplicación de modelo de simulación basada en agentes a gran escala para la ciudad de Santiago*. University of Chile, Carabineros de Chile. (2017). SIEC-2. In.
- CONASET. (2017). Estadísticas Generales. Retrieved from https://www.conaset.cl/wp-content/uploads/2017/07/Zonadeocurrencia2000_2016.xlsx
- Directorio de Transporte Público Metropolitano. (2013). *Informe de gestión*. Retrieved from [https://www.dtpm.cl/archivos/Memoria%202013-Web_Final%20\(1\).pdf](https://www.dtpm.cl/archivos/Memoria%202013-Web_Final%20(1).pdf)
- Elvik, R. (2007). *State-of-the-art approaches to road accident black spot management and safety analysis of road networks*: Transportøkonomisk institutt.
- Gazmuri, P., Muñoz, J. C., Rizzi, L. I., Fresard, F., & Cumsille, S. (2006). Reducción de la mortalidad por accidentes del tránsito en Chile: 10 medidas prioritarias. In: Camino al Bicentenario. Doce propuestas para Chile. Pontificia Universidad Católica de Chile, Ministerio del Interior, Ministerio de Planificación, Senado de la República. Chile.
- Golob, T. F., & Recker, W. W. (2004). A method for relating type of crash to traffic flow characteristics on urban freeways. *Transportation Research Part A: Policy and Practice*, 38(1), 53-80.
- Gomes, S. V. (2013). The influence of the infrastructure characteristics in urban road accidents occurrence. *Accident Analysis & Prevention*, 60, 289-297.
- Greibe, P. (2003). Accident prediction models for urban roads. *Accident Analysis & Prevention*, 35(2), 273-285.
- Hauer, E. (1997). *Observational before/after studies in road safety. Estimating the effect of highway and traffic engineering measures on road safety*.
- Instituto Nacional de Estadística. (2013). *Transporte y Telecomunicaciones, Anuario Parque Vehicular*. Retrieved from <http://www.ine.cl/estadisticas/economicas/transporte-y-comunicaciones>
- Instituto Nacional de Estadística. (2017). *Demografía y vitales*. Retrieved from <http://www.ine.cl/estadisticas/demograficas-y-vitales>
- IRTAD. (2017). *Benchmarking de la seguridad vial en América Latina*. Retrieved from <https://www.itf-oecd.org/benchmarking-seguridad-vial-am%C3%A9rica-latina>
- Kickhofer, B., Hosse, D., Turnera, K., & Tirachini, A. (2016). *Creating an open MATSim scenario from open data: The case of Santiago de Chile*. Retrieved from
- Layfield, R. E., Summersgill, I., Hall, R., & Chatterjee, K. (1996). Accidents at urban priority crossroads and staggered junctions. *TRL REPORT 185*.
- Lindsey, J. K. (2000). *Applying generalized linear models*: Springer Science & Business Media.
- Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, 44(5), 291-305.

- Lord, D., Washington, S. P., & Ivan, J. N. (2005). Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis & Prevention*, 37(1), 35-46.
- Maher, M., & Mountain, L. (2009). The sensitivity of estimates of regression to the mean. *Accident Analysis and Prevention*, 41(4), 861-868. doi:10.1016/j.aap.2009.04.020
- Maher, M. J., & Summersgill, I. (1996). A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis & Prevention*, 28(3), 281-296.
- Maluenda Nazar, N. s. R. (2012). *Modelo de evaluación en red de proyectos de reconstrucción de establecimientos educacionales*. Santiago de Chile, Santiago de Chile.
- MINSAL. (2017). Defunciones por accidentes de tránsito. Retrieved from http://www.deis.cl/wp-content/2017/gobCL-sitios-1.0/assets/SerieDefunciones_2000_2015.html
- Ley 20.285: Sobre acceso a la información pública, (2008).
- Mountain, L., Fawaz, B., & Jarrett, D. (1996). Accident prediction models for roads with minor junctions. *Accident Analysis & Prevention*, 28(6), 695-707.
- Nelder, J., & Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A*, 135, 370-384. nelder370135j. r. *Statist. Soc A*, 1972.
- New Zealand Transport Agency. (2013). *Economic Evaluation Manual*.
- Nilsson, G. (2004). Traffic safety dimensions and the power model to describe the effect of speed on safety. *Bulletin-Lunds Tekniska Högskola, Inst för Teknik och Samhälle, Lunds Universitet*, 221.
- Ley 20.285: Sobre acceso a la información pública, (2008).
- Rizzi, L. (2001). Fotorradares y seguridad vial: un análisis empírico bayesiano. *Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile*.
- SECTRA. (2007). Análisis y definición de una metodología para la evaluación social de impactos de proyectos sobre la seguridad vial en rutas interurbanas. In C. I. EIRL (Ed.), *Felipe Masjuán Ingeniería EIRL: SECTRA*
- SECTRA. (2013). *Mediciones de aforos de tráfico y perfiles de carga en servicios troncales en el Gran Santiago*. Retrieved from <http://www.sectra.gob.cl/biblioteca/detalle1.asp?mfn=3090>
- SECTRA. (2014). *Actualización y recolección de información del sistema de transporte urbano, IX etapa: Encuesta Origen Destino Santiago 2012*. Retrieved from <http://www.sectra.gob.cl/biblioteca/detalle1.asp?mfn=3253>
- SECTRA. (2014). Metodología para la evaluación social de la reducción de accidentes urbanos. In A. C. Ltda (Ed.): SECTRA.
- Turner, S., & Nicholson, A. (1998). Intersection accident estimation: the role of intersection location and non-collision flows. *Accident Analysis & Prevention*, 30(4), 505-517.
- WHO, W. H. O. (2015). *Global Status Report on Road Safety*. Retrieved from http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/
- Yañez, M. F. (2011). Contexto Estadístico de la Accidentabilidad Vial en Chile. In: Nacional de Seguridad de Tránsito CONASET.
- Ye, Z., Zhang, Y., & Lord, D. (2013). Goodness-of-fit testing for accident models with low means. *Accident Analysis & Prevention*, 61, 78-86.

10 Anexos

10.1 Tablas Base SIEC2 y Descripción de los Campos

Nombre campo	Descripción
Año	Año de ocurrencia del accidente
Crr_registro	Identificador del accidente
Cod_clase_accidente	Código clase de accidente, 1: acc. tránsito, 2: acc. Ferroviario
Fec_registro	Fecha del accidente en formato dd-mm-aa
Num_hora	Hora del accidente en formato hh-mm
Cod_comuna	Código de comuna donde ocurrió accidente
Region	Código región donde ocurrió accidente: 1-13
Cod_causa	Código causa probable del accidente, se consideran 59 posibles causas
Cod_tip_accid	Código tipo de accidente, se consideran 25 tipos de accidentes
Tipoaccidente	Glosa descriptiva del tipo de accidente
Cod_urbrur	Código sector del sector donde ocurrió el accidente
Urba_rur	Glosa descriptiva del tipo de sector (urbano, interurbano o vía férrea)
Cod_ubicrela	Código ubicación relativa del accidente
Ubicación relativa	Glosa descriptiva de la ubicación relativa del accidente, se consideran 18 tipos de accidente
Cod_estatmosf	Código del estado atmosférico al momento del accidente, se consideran 6 estados posibles
Estadoatmosferico	Glosa descriptiva del estado atmosférico al momento del accidente
Cod_calzada	Código clase de calzada
Calzada	Glosa descriptiva de la clase de calzada (unidireccional, bidireccional)
Cod_tipcalzad	Código tipo de material calzada, se consideran 6 tipos de material
Tipocalzada	Glosa descriptiva del tipo de calzada
Cod_estcalzad	Código estado de la calzada
Estadocalzada	Glosa descriptiva estado de la calzada (bueno, regular, malo)
Cod_condcalzada	Código condición de la calzada, se consideran 9 condiciones
Condicioncalzada	Glosa descriptiva condición de la calzada (seco, húmedo, con barro, etc.)
Calle1	Nombre de primera calle de referencia
Calle2	Nombre de segunda calle de referencia
Altura	Numeración frente a la que ocurrió el accidente si éste se produjo en un tramo de vía
Ruta1	Nombre de primer camino de referencia para accidentes interurbanos
Ruta 2	Nombre de segundo camino de referencia para accidentes interurbanos
Kilometro	Kilómetro en el que ocurrió el accidente si éste se produjo en un tramo de vía
Km_tren	Kilómetro de la vía férrea en el que ocurrió el accidente

Tabla 10-1: Estructura Tabla Accidentes-Base Siec-2.

Fuente: Análisis y Definición de una Metodología para la Evaluación Social de Impactos de Proyectos sobre la Seguridad Vial en Rutas Interurbanas (SECTRA, 2007).

Nombre campo	Descripción
Año	Año de ocurrencia accidente
Crr_registro	Identificador del accidente
Cod_calidad	Código para participación de la persona
Calidad	Glosa descriptiva para participación de la persona: conductor, pasajero o peatón
Codsexo	Código sexo persona, 1: masculino, 2: femenino
Num_edad	Edad de la persona
Cod_resultado	Código consecuencias accidente en la persona (1: muerto, 2: herido grave, 3: herido menos grave, 4: herido leve, 5: ileso)
Cod_cinturon	Código uso cinturón de seguridad o casco (0: se ignora, 1: sí, 2: no)
Cod_licencia	Código tipo de licencia de conducir (si es conductor)
Cod_condfisica	Código condición física (si es conductor)
Cond fisica	Código condición física (si es conductor)
Cod_nacionalidad	Código nacionalidad (si es conductor)
Nacionalidad	Glosa descriptiva de nacionalidad

Tabla 10-2: Estructura Tabla Personas-Base Siec-2.

Fuente: Análisis y Definición de una Metodología para la Evaluación Social de Impactos de Proyectos sobre la Seguridad Vial en Rutas Interurbanas (SECTRA, 2007).

Nombre campo	Descripción
Año	Año de ocurrencia accidente
Crr_registro	Identificador del accidente
Cod_vehiculo	Código tipo de vehículo, se consideran 30 tipos
Tg_vehiculos_idf_glosa	Glosa descriptiva del tipo de vehículo
Cod_servicio	Código tipo de servicio que realizaba el vehículo, se consideran 17 tipos de servicio
Tg_servicios_idf_glosa	Glosa descriptiva del tipo de servicio
Cod_direccion	Código dirección de viaje del vehículo, se considera 5 direcciones
Tg_direccion_idf_glosa	Glosa descriptiva dirección de viaje
Cod_consecuencia	Código consecuencias en el vehículo, se considera 3 opciones
Tg_consecuencia_idf_glosa	Glosa descriptiva consecuencias en el vehículo
Cod_maniobra	Código de maniobra realizada por el vehículo, se considera 18 opciones de maniobras
Tg_maniobras_idf_glosa	Glosa descriptiva maniobra del vehículo

Tabla 10-3: Estructura Tabla Vehículos-Base Siec-2.

Fuente: Análisis y Definición de una Metodología para la Evaluación Social de Impactos de Proyectos sobre la Seguridad Vial en Rutas Interurbanas (SECTRA, 2007).

10.2 Distribuciones de accidentes geolocalizados por año

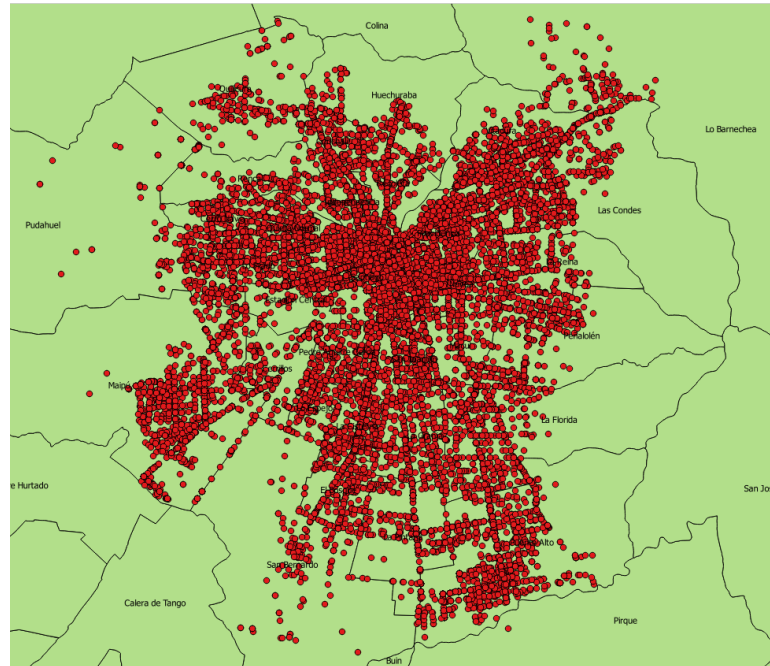


Figura 16: Distribución de accidentes en la región metropolitana para el año 2011.
Fuente: Elaboración propia en base a datos SIEC-2 y BCN.

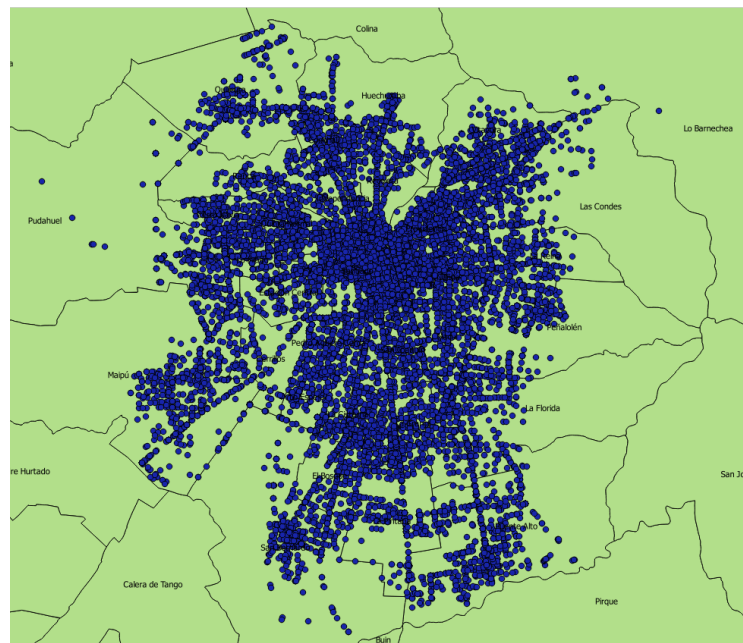


Figura 17: Distribución de accidentes en la región metropolitana para el año 2012.
Fuente: Elaboración propia en base a datos SIEC-2 y BCN.

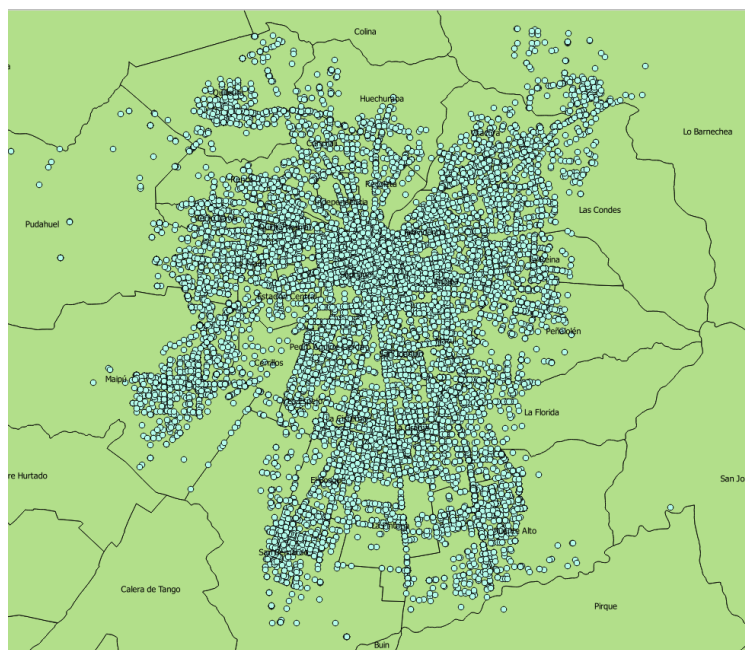


Figura 18: Distribución de accidentes en la región metropolitana para el año 2013.
Fuente: Elaboración propia en base a datos SIEC-2 y BCN.

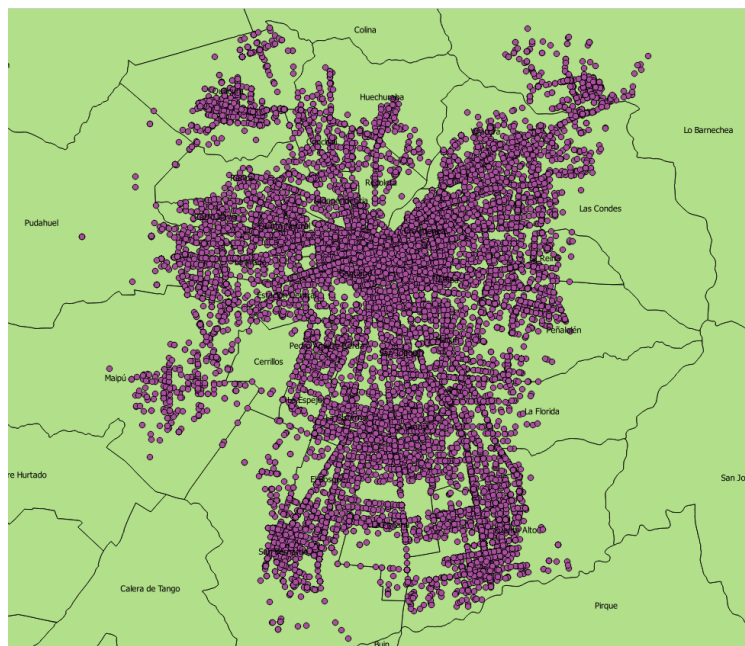
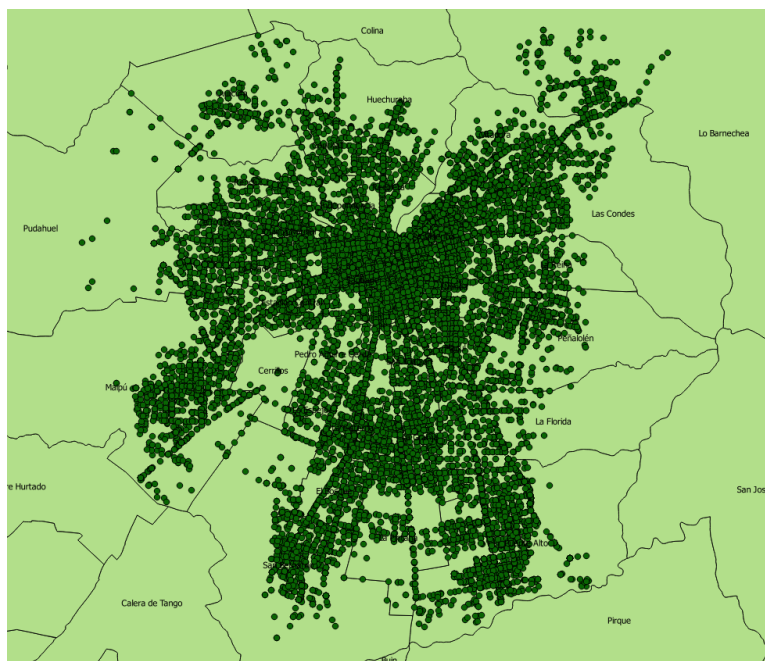


Figura 19: Distribución de accidentes en la región metropolitana para el año 2014.
Fuente: Elaboración propia en base a datos SIEC-2 y BCN.



*Figura 20: Distribución de accidentes en la región metropolitana para el año 2015.
Fuente: Elaboración propia en base a datos SIEC-2 y BCN.*

10.3 Descripción de la Base de Datos de Bienes raíces del SII

INFORMACIÓN BÁSICA ROLES NO AGRÍCOLAS	
Campo	Descripción
0	Año y Semestre
1	Código SII de la Comuna
2	Número de Manzana
3	Número de Predial
4	Dirección o nombre del predio
5	Avalúo fiscal total
6	Contribución semestral (con aseo)
7	Código de destino principal
8	Avalúo exento de la propiedad
9	Código SII de la Comuna Rol Bien Común 1
10	Número de Manzana Rol Bien Común 1
11	Número de predio Rol Bien Común 1
12	Código SII de la Comuna Rol Bien Común 2
13	Número de Manzana Rol Bien Común 2
14	Número de predio Rol Bien Común 2
15	Superficie total del terreno de la propiedad (sin decimales, en m ²)

Tabla 10-4: Campos de la Base de Datos de Bienes Raíces BBRR para la Información Básica de Roles No Agrícolas.
Fuente: SII.

INFORMACIÓN DE TERRENOS Y CONSTRUCCIONES NO AGRÍCOLAS	
Campo	Descripción
0	Año y semestre
1	Código SII de la Comuna
2	Número de Manzana
3	Número de Predial
4	Número correlativo de la línea de construcción
5	Código del material estructural de la línea de construcción
6	Código de calidad de la línea de construcción
7	Año de la línea de construcción
8	Superficie de la línea de construcción (sin decimales, en m ² o m ³ según tipo de construcción)
9	Código de destino de la línea de construcción
10	Código de condición especial de la línea de construcción

Tabla 10-5: Campos de la Base de Datos de Bienes Raíces BBRR para la Información de Terrenos y Construcciones No Agrícolas.
Fuente: SII.

10.4 Algoritmo de rutas mínimas para datos OD de la EOD 2012

10.4.1 Generación de la Red

- 1) Se genera una nueva base de datos en PostgreSQL
- 2) Se genera la extensión de PostGIS y Pgrouting
- 3) Se importa el archivo Shapefile con la red
- 4) Se agregan las columnas "source" y "target"
- 5) Se genera la consulta:

```
SELECT pgr_createTopology('red_memoria',0.0001,'the_geom','gid')
```

- 6) Se importan los pares orígenes destinos del desarrollo explicado a continuación.

10.4.2 Generación de Tablas con Nodos Origen y Destino

1. Filtro de las etapas por modo para los datos que puedan ser georreferenciados y ser útiles en el análisis de rutas mínimas
2. Creación de shapefile con los orígenes de las etapas. Lo mismo para los destinos.
 - 2.1. Coordenadas EPSG: 32719 utm zona 19 S.
3. Obtener los vértices de los nodos de la red con la conexión a PostGIS que posee QGIS al conectarse con la red ya incluida en Postgres.
4. Con estos archivos obtener la tabla de cercanía (Herramienta "Generate Near Table") con el software ArcGis.
 - 4.1. Se realiza para los 2000 primeros nodos más cercanos.
5. Se guarda la tabla para luego importarla al software Office Access.
 - 5.1. Se importa:
 - 5.1.1. Tabla de cercanía de los orígenes con los vértices.
 - 5.1.2. Tabla de cercanía de los destinos con los vértices.
 - 5.1.3. Vértices con los campos de la zonificación EOD.
 - 5.1.4. BD orígenes y destinos
6. Se genera una consulta en la que se devuelve toda la información de los orígenes o destinos que cumpla el criterio de que la zona del vértice sea igual a la zona declarada en las tablas O/D. Se agrega el campo de distancia para que entregue la menor distancia pues es ese el nodo más cercano que cumple el criterio de que la zona de origen o destino declarada sea efectivamente la del nodo cercano.
7. Se exporta esa tabla en formato CSV (Comma-separated values) para luego ser importada a la base de datos de la red en Postgres.
8. Con esa tabla se genera la siguiente consulta que entrega como resultado las rutas mínimas mediante el algoritmo Dijkstra implementado en la extensión PostGIS:

```
SELECT source, target, pgr_dijkstra('SELECT gid As id,
source, target, length As cost FROM
red_memoria',source,target,false,false) As paths
FROM public."ODcam";
```

9. Con la tabla de resultado de la consulta anterior, se exporta para ser trabajada en Office Access, donde se genera la relación con la tabla de arcos y se obtienen las zonas por donde pasan las rutas, mediante una consulta de conteo.

10.5 Artículos Científicos con resumen de variables explicadas y explicativas

Variables y descripción		Referencia								
		(M. J. Maher & Summersgill, 1996)	(Golob & Recker, 2004)	(Mountain et al., 1996)	(Greibe, 2003)	(Ye et al., 2013)	(Elvik, 2007)	(Gomes, 2013)	(New Zealand Transport Agency, 2013)	(AASHTO, 2010)
Variables explicadas	Número de accidentes			Sí		Sí			Sí	
	Número de accidentes en tramos de vía	Sí		Sí	Sí		Sí	Sí	Sí	Sí
	Número de accidentes según gravedad de lesiones		Sí					Sí		
	Número de accidentes según tipo de usuario	Sí								Sí
	Número de accidentes por tipo de colisión o cantidad de vehículos involucrados		Sí							Sí
Variables explicativas	Variables operativas	Flujo total				Sí	Sí	Sí	Sí	Sí
		Flujos por tipo de usuario	Sí							
		Flujos por pista		Sí						
		Flujo en arco			Sí			Sí		
		Flujo peatonal	Sí							Sí
		Otra desagregación/agregación de flujo		Sí						
		Condición climática al momento del accidente		Sí						
	Variables de infraestructura	Longitud del arco/segmento	Sí		Sí		Sí	Sí	Sí	Sí
		Número de intersecciones en el segmento de vía			Sí					
		Ancho de pista/tramo				Sí		Sí		
		Número de pistas						Sí	Sí	Sí
		Condición de luminosidad o visibilidad al momento del accidente		Sí						Sí
Variables explicativas	Variables de infraestructura	Número de accesos, salidas, caminos secundarios o cantidad de intersecciones en el segmento de vía				Sí		Sí		Sí
		Presencia de mediana								Sí
		Presencia de pistas de viraje								Sí
		Presencia de objetos fijos al costado de la vía								Sí
	Variables de regulación de la circulación	Velocidad límite				Sí		Sí		
		Uso de suelo aledaño				Sí				
		Regulaciones de estacionamiento, ocupación y actividad				Sí				Sí
		Tipo o clasificación de vía urbana						Sí		

Tabla 10-6: Resumen de variables explicadas y explicativas en MPA: caso tramos de vía.
Fuente: Metodología para la Evaluación Social de la Reducción de Accidentes Urbanos (SECTRA, 2014).

Variables y descripción		Referencia										
		(M. J. Maher & Summersgill, 1996)	(Layfield et al., 1996)	(Mountain et al., 1996)	(Greibe, 2003)	(Ye et al., 2013)	(Turner & Nicholson, 1998)	(Gomes, 2013)	(New Zealand Transport Agency, 2013)	(AASHTO, 2010)		
Variables explicadas	Número de accidentes		Si	Si		Si			Si			
	Número de accidentes en intersección	Si	Si	Si	Si		Si	Si	Si	Si		
	Número de accidentes según nivel de lesiones							Si				
	Número de accidentes por tipo de usuario	Si	Si							Si		
	Número de accidentes por tipo de movimiento vehicular	Si	Si				Si					
	Número de accidentes por tipo de colisión o cantidad de vehículos involucrados			Si							Si	
Variables explicativas	Variables de operación	Flujo total	Si	Si				Si	Si			
		Flujos en vías prioritaria y secundaria	Si	Si		Si	Si			Si	Si	
	Flujos de vehículos según tipo de movimiento	Si	Si				Si			Si		
	Flujos por tipo de usuario		Si					Si		Si		
	Flujos por pista		Si							Si		
	Flujo en arco		Si	Si								
	Flujo peatonal	Si	Si					Si				
	Proporción de vehículos de 2 ruedas.	Si										
	Vehículos por unidad de tiempo de verde	Si										
	Tiempo de entreverde	Si										
	Presencia de fase de viraje a la izquierda (si está señalizado)										Si	
	Velocidad de vehículo/de flujo libre/de diseño		Si							Si		
	Variables explicativas	Variables de operación	Longitud de cola		Si							
Número de ramas con tráfico en una sola dirección								Si				
Presencia/número de pistas de viraje									Si		Si	
Presencia de isla central.												
Variables de Infraestructura		Ancho de pista/tramo		Si					Si			
		Número de pistas/ramas		Si		Si					Si	
		Presencia de cruce peatonal		Si								
		Ángulo de viraje	Si	Si								
		Desplazamiento de la línea central con respecto a su continuación en la intersección.	Si									
		Presencia de isla central.	Si	Si								
		Condición de luminosidad o visibilidad al momento del accidente		Si						Si	Si	
		Presencia de paraderos		Si								
		Balace de pista							Si			
		Señalizaciones o regulación del cruce	Si	Si								Si
Variables de regulación de la circulación		Demarcaciones		Si								
		Uso de suelo aledaño		Si								
		Regulaciones de estacionamiento, ocupación y actividad		Si								
		Presencia de cámaras de semáforo en rojo										Si
		Velocidad límite		Si								
Otras Variables		Sexo y edad conductores		Si								
	Early cut-off or late release.	Si										
	Si la intersección estaba en la capital o no		Si									

Tabla 10-7: Resumen de variables explicadas y explicativas en MPA: caso intersecciones.
Fuente: Metodología para la Evaluación Social de la Reducción de Accidentes Urbanos (SECTRA, 2014).

Número Documento	Variables Explicadas (por unidad de tiempo)		Variables Explicativas								
	Cant. de Acc. Totales (incluyendo peatones)	Cant. De Acc con lesionados excluyendo peatones	Variables de Operación						Variables de Infraestructura		
			Flujo total/solo vehicular	Flujos de vehículos según tipo de	Flujo por tipo de usuarios	Flujo en arco/circulando la rotonda	Flujo peatonal	Velocidad de vehículo/de flujo	Número de pistas/salidas	Condición de luminosidad o visibilidad	
(Gomes, 2013)	Si	Si	Si							Si	
(New Zealand Transport Agency, 2013)	Si		Si	Si	Si	Si	Si	Si	Si		Si

Tabla 10-8: Modelos de Rotonda.
Fuente: Elaboración propia.

Número Documento	Variables Explicadas (por unidad de tiempo)		Variables Explicativas		
	Cant. de Acc. Totales en intersección	Cant. De Acc. por tipo de movimiento	Variables de Operación		
			Flujo total	Flujos por los que llegan en la vía mayor, vía menor	Flujos de vehículos según tipo de
(Turner & Nicholson, 1998)	Si	Si	Si	Si	Si

Tabla 10-9: Modelo de Red.
Fuente: Elaboración propia.

10.6 Estadísticos descriptivos de las variables

Variables	Unidades de la variable	Rango	Máximo	Suma	Media		Desviación estándar	Varianza	Asimetría		Curtosis	
		Estadístico	Estadístico	Estadístico	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico	Error estándar	Estadístico	Error estándar
prom acc tot	Acc/año	156,00	156,00	17.368,25	23,60	0,71	19,34	373,96	2,02	0,09	6,66	0,18
prom lesionados tot	Lesionados/año	83,75	83,75	10.556,25	14,34	0,45	12,24	149,92	1,77	0,09	4,35	0,18
prom fallecidos tot	Fallecidos/año	2,50	2,50	225,50	0,31	0,01	0,36	0,13	1,77	0,09	4,31	0,18
prom acc no atrop	Acc/año	140,75	140,75	15.226,50	20,69	0,63	16,96	287,52	2,00	0,09	6,65	0,18
prom lesionados no atrop	Lesionados/año	69,75	69,75	8.270,50	11,24	0,36	9,86	97,27	1,87	0,09	5,19	0,18
prom fallecidos no atrop	Fallecidos/año	1,50	1,50	114,50	0,16	0,01	0,24	0,06	1,88	0,09	4,01	0,18
prom acc atrop	Acc/año	24,00	24,00	2.141,75	2,91	0,11	2,91	8,46	2,40	0,09	8,98	0,18
prom lesionados atrop	Lesionados/año	24,50	24,50	2.285,75	3,11	0,11	3,11	9,69	2,28	0,09	7,63	0,18
prom fallecidos atrop	Fallecidos/año	2,00	2,00	111,00	0,15	0,01	0,25	0,06	2,70	0,09	10,42	0,18
DummyStgo	-	1,00	1,00	52,00	0,07	0,01	0,26	0,07	3,36	0,09	9,30	0,18
DummyCerrillos	-	1,00	1,00	16,00	0,02	0,01	0,15	0,02	6,57	0,09	41,31	0,18
DummyCerroNavia	-	1,00	1,00	14,00	0,02	0,01	0,14	0,02	7,06	0,09	47,92	0,18
DummyConchalí	-	1,00	1,00	20,00	0,03	0,01	0,16	0,03	5,83	0,09	32,05	0,18
DummyElBosque	-	1,00	1,00	17,00	0,02	0,01	0,15	0,02	6,36	0,09	38,59	0,18
DummyEstación Central	-	1,00	1,00	23,00	0,03	0,01	0,17	0,03	5,40	0,09	27,23	0,18
DummyHuechuraba	-	1,00	1,00	14,00	0,02	0,01	0,14	0,02	7,06	0,09	47,92	0,18
DummyIndependencia	-	1,00	1,00	15,00	0,02	0,01	0,14	0,02	6,80	0,09	44,40	0,18
DummyLaCisterna	-	1,00	1,00	17,00	0,02	0,01	0,15	0,02	6,36	0,09	38,59	0,18
DummyLaFlorida	-	1,00	1,00	45,00	0,06	0,01	0,24	0,06	3,67	0,09	11,51	0,18
DummyLaGranja	-	1,00	1,00	16,00	0,02	0,01	0,15	0,02	6,57	0,09	41,31	0,18
DummyLaPintana	-	1,00	1,00	15,00	0,02	0,01	0,14	0,02	6,80	0,09	44,40	0,18
DummyLaReina	-	1,00	1,00	22,00	0,03	0,01	0,17	0,03	5,53	0,09	28,69	0,18
DummyLasCondes	-	1,00	1,00	37,00	0,05	0,01	0,22	0,05	4,12	0,09	15,06	0,18
DummyLoBarrilete	-	1,00	1,00	13,00	0,02	0,00	0,13	0,02	7,34	0,09	51,99	0,18
DummyLoEspejo	-	1,00	1,00	11,00	0,01	0,00	0,12	0,01	8,01	0,09	62,36	0,18
DummyLoPrado	-	1,00	1,00	12,00	0,02	0,00	0,13	0,02	7,65	0,09	56,74	0,18
DummyMacul	-	1,00	1,00	18,00	0,02	0,01	0,15	0,02	6,17	0,09	36,17	0,18
DummyMaipo	-	1,00	1,00	41,00	0,06	0,01	0,23	0,05	3,88	0,09	13,11	0,18
DummyÑuñoa	-	1,00	1,00	30,00	0,04	0,01	0,20	0,04	4,65	0,09	19,72	0,18
DummyPAC	-	1,00	1,00	17,00	0,02	0,01	0,15	0,02	6,36	0,09	38,59	0,18
DummyPeñalolen	-	1,00	1,00	28,00	0,04	0,01	0,19	0,04	4,84	0,09	21,48	0,18
DummyProvidencia	-	1,00	1,00	23,00	0,03	0,01	0,17	0,03	5,40	0,09	27,23	0,18
DummyPudahuel	-	1,00	1,00	25,00	0,03	0,01	0,18	0,03	5,16	0,09	24,65	0,18
DummyQtaNormal	-	1,00	1,00	21,00	0,03	0,01	0,17	0,03	5,68	0,09	30,29	0,18
DummyRecoleta	-	1,00	1,00	24,00	0,03	0,01	0,18	0,03	5,27	0,09	25,88	0,18
DummyRenca	-	1,00	1,00	17,00	0,02	0,01	0,15	0,02	6,36	0,09	38,59	0,18
DummySanJoaquín	-	1,00	1,00	20,00	0,03	0,01	0,16	0,03	5,83	0,09	32,05	0,18
DummySanMiguel	-	1,00	1,00	14,00	0,02	0,01	0,14	0,02	7,06	0,09	47,92	0,18
DummySanRamón	-	1,00	1,00	13,00	0,02	0,00	0,13	0,02	7,34	0,09	51,99	0,18
DummyVitacura	-	1,00	1,00	19,00	0,03	0,01	0,16	0,03	5,99	0,09	34,00	0,18
DummyPuente Alto	-	1,00	1,00	42,00	0,06	0,01	0,23	0,05	3,83	0,09	12,68	0,18
DummySanBernardo	-	1,00	1,00	25,00	0,03	0,01	0,18	0,03	5,16	0,09	24,65	0,18

Variables	Unidades de la variable	Rango	Máximo	Suma	Media		Desviación estándar	Varianza	Asimetría		Curtosis	
		Estadístico	Estadístico	Estadístico	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico	Error estándar	Estadístico	Error estándar
Área zona	Km²	998,44	998,81	366.146,97	497,48	10,52	285,32	81.407,04	0,04	0,09	-1,20	0,18
Vialidad sin autopistas	km	107,21	107,21	12.164,74	16,53	0,47	12,86	165,28	2,91	0,09	12,16	0,18
Vialidad total	Km	101,42	101,42	11.792,11	16,02	0,46	12,38	153,16	2,86	0,09	12,01	0,18
Total de predios	Predios	65.934,00	65.934,00	2.993.629,00	4.067,43	193,16	5.240,21	27.459.750,07	4,37	0,09	33,31	0,18
COMERCIO	Predios por uso de suelo	3.823,00	3.823,00	87.099,00	118,34	8,86	240,44	57.810,89	7,50	0,09	87,08	0,18
DEPORTE Y RECREACION	Predios por uso de suelo	132,00	132,00	2.033,00	2,76	0,26	7,10	50,44	9,90	0,09	155,22	0,18
EDUCACION CULTURA	Predios por uso de suelo	145,00	145,00	8.901,00	12,09	0,61	16,48	271,52	3,56	0,09	17,75	0,18
HOTEL MOTEL	Predios por uso de suelo	662,00	662,00	1.873,00	2,54	1,02	27,71	767,76	20,24	0,09	453,39	0,18
HABITACIONAL	Predios por uso de suelo	22.425,00	22.425,00	2.144.285,00	2.913,43	114,26	3.099,82	9.608.871,93	2,47	0,09	8,03	0,18
INDUSTRIA	Predios por uso de suelo	690,00	690,00	18.470,00	25,10	2,26	61,30	3.757,82	5,88	0,09	46,61	0,18
BIENES COMUNES	Predios por uso de suelo	470,00	470,00	20.874,00	28,36	2,14	58,15	3.381,97	3,98	0,09	20,19	0,18
BODEGA Y ALMACENAJE	Predios por uso de suelo	10.605,00	10.605,00	240.103,00	326,23	36,20	982,00	964.317,09	5,27	0,09	35,29	0,18
MINERIA	Predios por uso de suelo	21,00	21,00	71,00	0,10	0,04	1,11	1,24	15,15	0,09	245,98	0,18
OFICINA	Predios por uso de suelo	6.381,00	6.381,00	51.411,00	69,85	13,99	379,61	144.102,09	10,72	0,09	139,73	0,18
ADMINISTRACION PUBLICA Y DEFENSA	Predios por uso de suelo	33,00	33,00	715,00	0,97	0,12	3,16	9,99	5,85	0,09	41,82	0,18
CULTO	Predios por uso de suelo	73,00	73,00	3.543,00	4,81	0,25	6,70	44,93	4,42	0,09	30,93	0,18
SALUD	Predios por uso de suelo	117,00	117,00	1.838,00	2,50	0,28	7,63	58,24	8,13	0,09	90,90	0,18
TRANSPORTE Y TELECOMUNICACIONES	Predios por uso de suelo	39,00	39,00	933,00	1,27	0,12	3,13	9,82	5,09	0,09	38,62	0,18
OTROS NO CONSIDERADOS	Predios por uso de suelo	187,00	187,00	6.027,00	8,19	0,57	15,47	239,42	5,75	0,09	45,93	0,18
SITIO ERIAZO	Predios por uso de suelo	1.684,00	1.684,00	40.601,00	55,16	5,59	151,65	22.996,96	5,98	0,09	44,45	0,18
ESTACIONAMIENTO	Predios por uso de suelo	35.194,00	35.194,00	364.852,00	495,72	71,58	1.941,92	3.771.058,08	9,97	0,09	148,20	0,18
Suma de Avalúo Fiscal de los predios	\$	4.833.657.145,952,00	4.833.657.145.952,00	124.273.530.998.637,00	168.849.906.248,15	13.172.942,364,71	357.372.967.864,54	127.715.438.160.309.000.000,00	6,24	0,09	56,82	0,18
Flujo_vehicular	Vehiculos/zona	179.454,00	179.454,00	15.253.804,00	20.725,28	884,84	24.005,10	576.245,003,46	2,58	0,09	8,86	0,18
Transbordos	Transbordo s/zona	93.569,82	93.569,82	2.005.794,99	2.725,26	288,12	7.816,48	61.097.354,60	5,56	0,09	40,83	0,18
Carga_bici	Bicicletas/zona	17.847,53	17.847,53	2.070.101,58	2.812,64	110,29	2.992,09	8.952.629,36	2,28	0,09	6,27	0,18
Carga_cam	Peatones/zona	71.882,36	71.882,36	6.346.813,02	8.623,39	298,35	8.094,16	65.515.373,31	2,75	0,09	12,33	0,18
Carga_moto	Motocicleta s/zona	9.633,07	9.633,07	822.112,69	1.117,00	50,56	1.371,75	1.881.702,02	1,94	0,09	4,51	0,18
Flujo_buses_ts	Buses/zona	5.022,08	5.022,08	883.938,88	1.201,00	27,45	744,70	554.582,53	1,51	0,09	3,70	0,18
flujos_auto_per s	Persona en vehículo/zona	253.030,14	253.030,14	21.507.863,64	29.222,64	1.247,62	33.847,20	1.145.632.691,37	2,58	0,09	8,86	0,18
flujos_bus_per s	Persona en bus/zona	128.840,77	128.840,77	22.677.336,99	30.811,60	704,23	19.105,26	365.010.887,12	1,51	0,09	3,70	0,18

Tabla 10-10: Estadísticos descriptivos de las variables consideradas.
Fuente: Elaboración propia.

10.7 Pestañas de la modelación lineal generalizada en SPSS

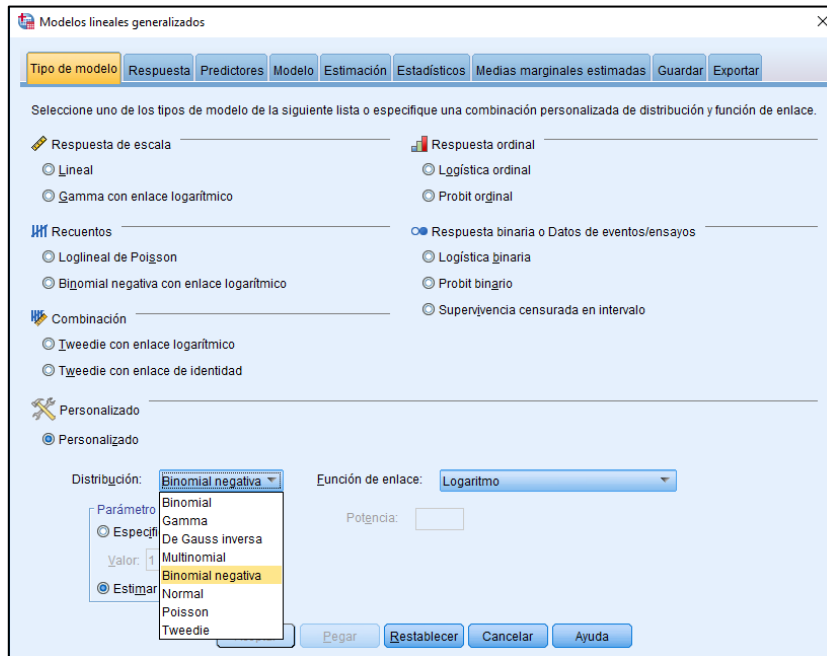


Figura 21: Pestaña Tipo de modelo en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

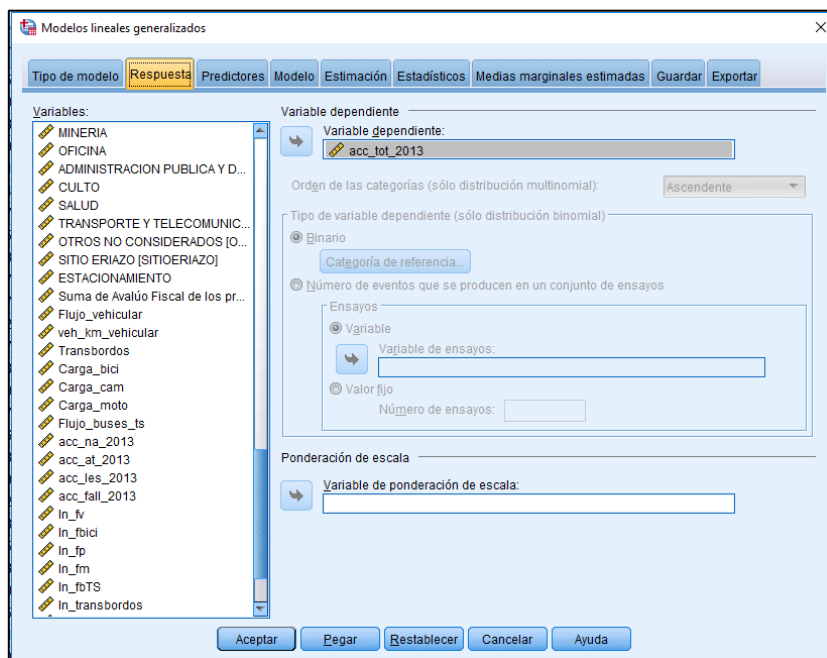


Figura 22: Pestaña de Respuesta en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

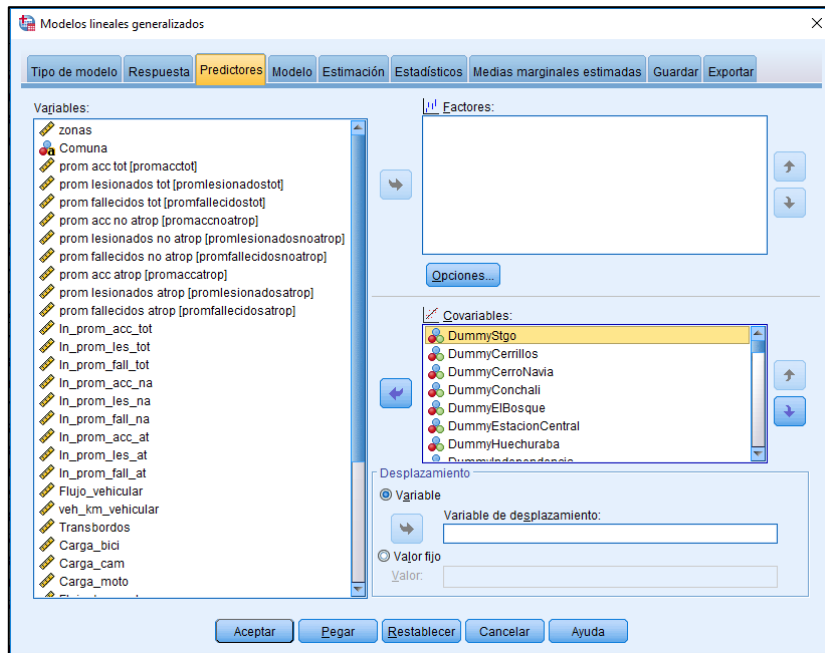


Figura 23: Pestaña Predictores en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

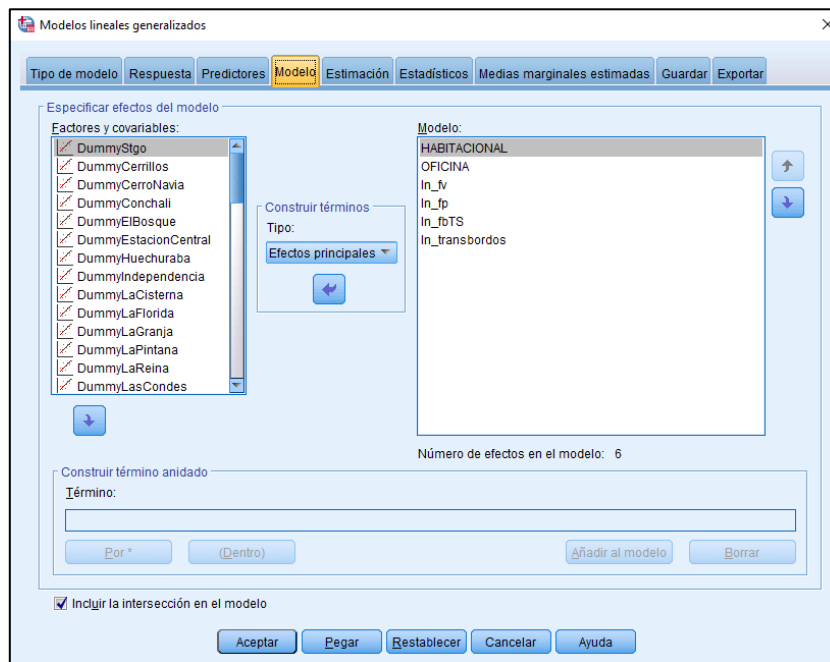


Figura 24: Pestaña de Modelo en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

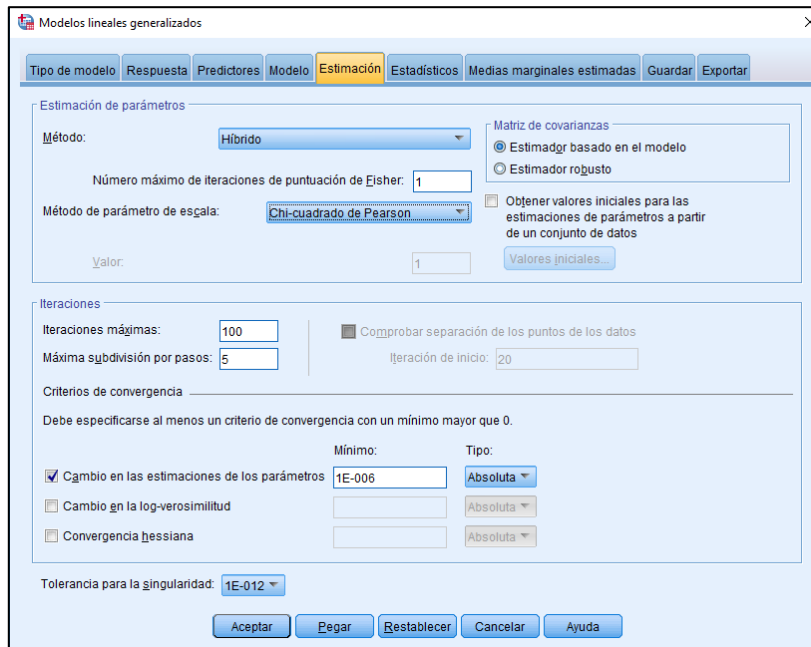


Figura 25: Pestaña de Estimación en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

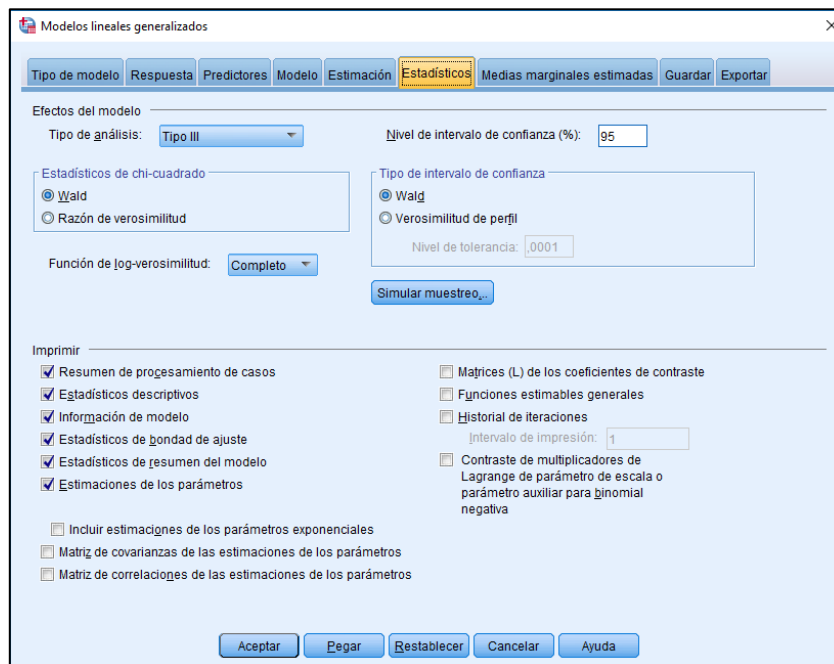


Figura 26: Pestaña Estadísticos en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

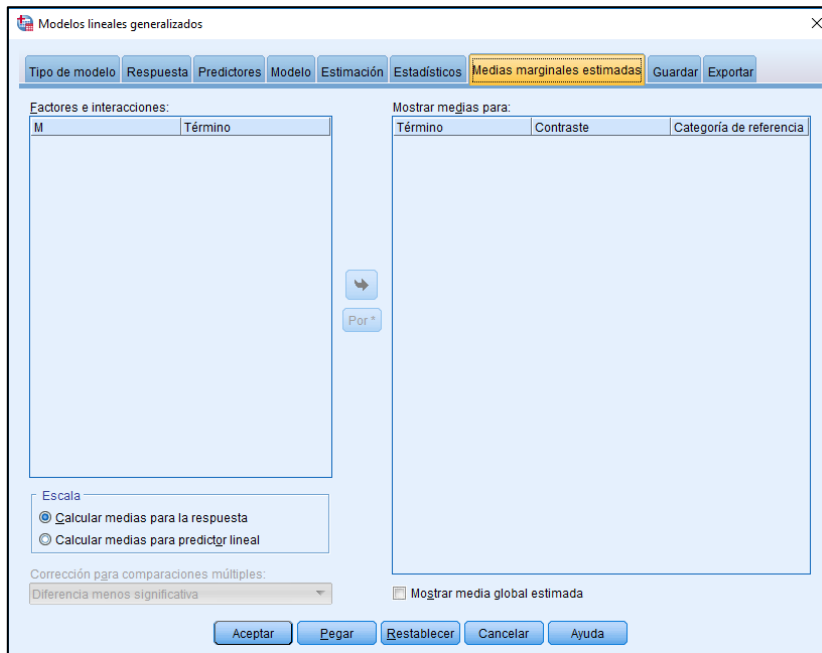


Figura 27: Pestaña Medias marginales estimadas en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

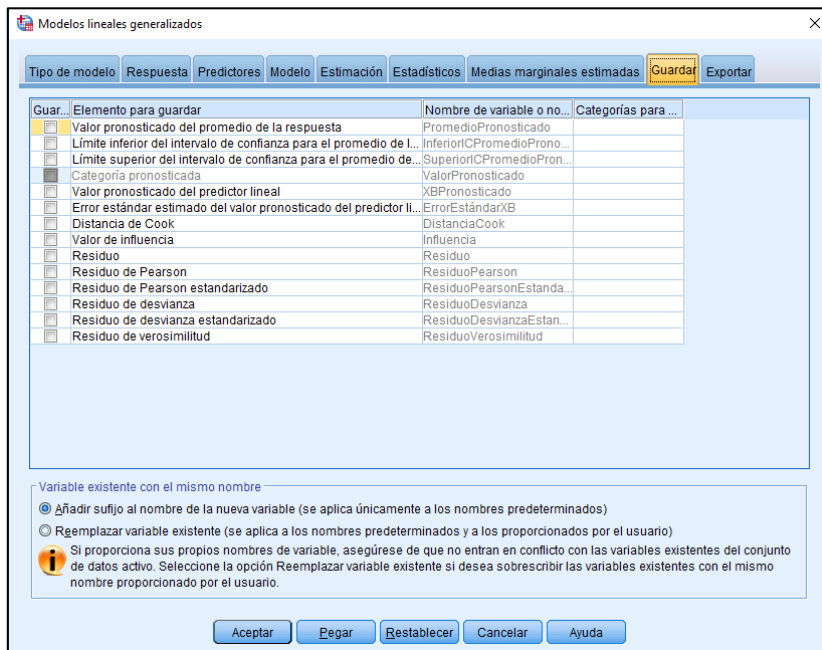


Figura 28: Pestaña Guardar en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.

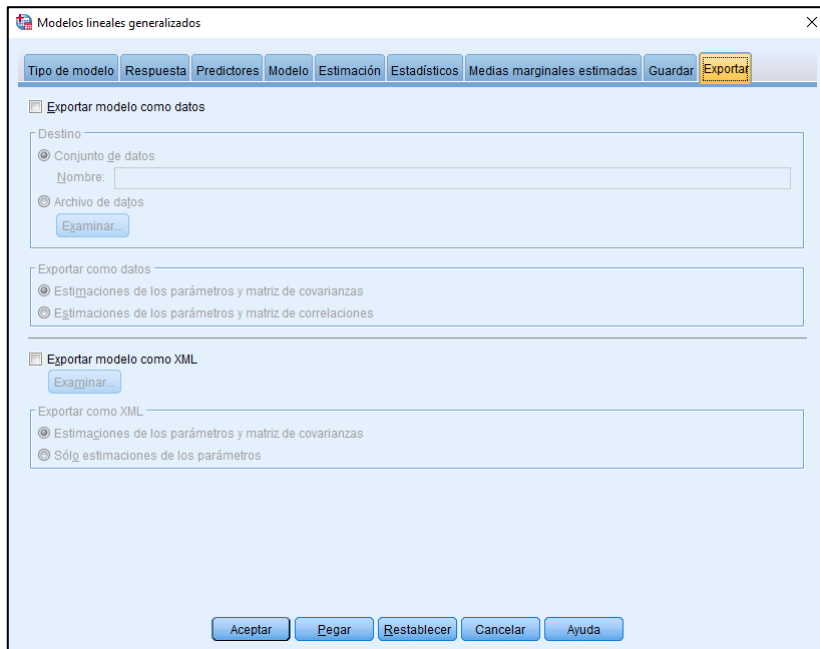


Figura 29: Pestaña Exportar en SPSS para la modelación lineal generalizada.
Fuente: Elaboración propia en base a SPSS.