



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

ESTIMACIÓN DE PROBABILIDADES DE ACCIDENTES BASADA EN ESTADOS DE TRÁFICO EN AUTOPISTAS URBANAS

CRISTIAN NICOLÁS ZÚÑIGA GONZÁLEZ

Tesis para optar al grado de
Magíster en Ciencias de la Ingeniería

Profesor Supervisor:
JUAN CARLOS HERRERA MALDONADO

Santiago de Chile, Julio, 2014.

© 2014, Zúñiga G., Cristian.



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

ESTIMACIÓN DE PROBABILIDADES DE ACCIDENTES BASADA EN ESTADOS DE TRÁFICO EN AUTOPISTAS URBANAS

CRISTIAN NICOLÁS ZÚÑIGA GONZÁLEZ

Tesis presentada a la Comisión integrada por los profesores:

JUAN CARLOS HERRERA MALDONADO

LUIS IGNACIO RIZZI CAMPANELLA

ALEJANDRO MAXIMILIANO TUDELA ROMÁN

ENZO ENRIQUE SAUMA SANTIS

Para completar las exigencias del grado de
Magíster en Ciencias de la Ingeniería

Santiago de Chile, Julio, 2014

Dedicatoria.

*“A la U, mi familia, amigos y
en especial a mi querido abuelo
Edgardo”*

AGRADECIMIENTOS

En primer lugar, quisiera agradecer a Dios. Sin Él esto no hubiese sido posible, con su ayuda y fuerza pude sacar adelante esta investigación que me otorga el grado de Magister en Ciencias de la Ingeniería.

En segundo lugar, agradecer a mi familia que me ofreció todo el apoyo en momentos de estrés. En especial, a mi querida madre que tuvo que aguantar esos momentos de frustración que sentí, porque las cosas no se daban como yo quería. A mi padre por ser el gran patrocinador de este proceso: “Viejo eres un crack/flaite, mil gracias!”. También, a mis hermanos Felipe y Matías que siempre estuvieron dándome buenos consejos y aporte en ideas cuando tuve dudas. Felipe en especial le agradezco por su *feedback*, ya que entiende a la perfección el proceso que viví en este postgrado. No puedo dejar de lado al Bosquiño, siempre un perro que te brinda cariño y juego en los mejores momentos.

En tercer lugar, a mis amigos y compañeros de postgrado. Gabriel, Chris, Andrés, Nico, Mati, Cami, Nacho, Natan, Niko, Francisco, Cesas, Felipe y muchos otros (si alguien se me queda afuera, tiene derecho a un bebestible). Sin su ayuda, consejos, sugerencias e ideas, esto no pudo haberse hecho realidad.

En cuarto lugar, no puedo dejar de agradecer a mis profesores. Juanca y Luis siempre estuvieron ahí para atender mis dudas y sugerencias, con el fin de realizar una buena e interesante investigación.

Finalmente, agradecer a mis grandes amigos y amigas que están conmigo en las buenas y en las malas. Lopez, Rata, Mosella, Pancho, Juanca, Pérez, Pimentel, Javi, Jo, Maida, Liss, Cata, Mono y Vicho (nuevamente, me disculpo si se me fue alguien).

A todos ustedes les digo: SALUD!

TABLA DE CONTENIDOS

	Pág.
AGRADECIMIENTOS	ii
RESUMEN	x
ABSTRACT	xi
1 INTRODUCCIÓN	1
1.1 Motivación	1
1.2 Objetivos e Hipótesis	3
1.3 Alcances	3
1.4 Contenido	4
2 MARCO TEÓRICO	5
2.1 Revisión Bibliográfica	6
2.1.1 Estudios de Predicción de Ocurrencia de Accidentes	6
2.1.2 Estudios de Identificación de <i>Hotspots</i>	9
2.1.3 Estudios de Severidad de Accidentes	11
2.2 Modelos de Predicción	14
2.2.1 Matched Case-Control Logistic Regression (MCCLR)	14
2.3 Modelos de Elección Discreta	16
2.3.1 Modelo Logit Multinomial (MNL)	17
2.3.2 Modelo Logit Jerárquico (HLM)	18
2.3.3 Modelo Logit Mixto (MLM)	19
2.4 Indicadores y <i>Tests</i> Estadísticos	20
2.4.1 Riesgo Relativo (<i>Relative Risk</i>)	20
2.4.2 Porcentaje Correctamente Predicho o FPR (<i>First Preference Recovery</i>)	21
2.4.3 Test de Razón de Verosimilitud	21
3 DISEÑO EXPERIMENTAL	22
3.1 Descripción General de Autopista Central	23
3.2 Definición de observaciones	24

3.3	Base de Datos.....	25
3.3.1	Accidentes.....	25
3.3.2	Información de Espiras	26
3.3.3	Información de Geometría e Infraestructura.....	29
3.3.4	Información de Condiciones Climáticas.....	30
3.3.5	Estados de Tráfico y Variables Explicativas.....	31
3.3.6	Caso-Control	38
4	ESTIMACION DE MODELOS Y RESULTADOS	40
4.1	Logit Multinomial (MNL).....	41
4.1.1	Funciones de Propensión	42
4.1.2	Signos Esperados	45
4.1.3	Resultados Generales de Estimación	47
4.1.4	Parámetros Estimados del Modelo Binario.....	49
4.1.5	Parámetros Estimados del Modelo de 3 Alternativas DP	53
4.1.6	Parámetros Estimados del Modelo de 3 Alternativas DT	57
4.1.7	Parámetros Estimados del Modelo de 5 Alternativas	61
4.1.8	Análisis Comparativo de Modelos.....	66
4.1.9	Validación de los Modelos.....	70
4.2	Logit Mixto (MLM)	72
4.2.1	Cálculo de Ponderadores.....	73
4.2.2	Especificación de los MLM	74
4.2.3	Comparación y resultados de modelos.....	75
4.3	Análisis e Interpretación de Parámetros de Mejor Modelo.....	80
4.3.1	Significancia, Signo y Magnitud de los Parámetros	80
4.3.2	Parámetros asociados a interacciones	80
4.3.3	Parámetros asociados a Desviaciones Estándar de Velocidad.....	81
4.3.4	Parámetros asociados a Estados de Tráfico	82
4.3.5	Parámetros asociados a Variables Geométricas.....	83
4.3.6	Resto de los Parámetros	84

5	PROPUESTAS BASADAS EN LOS RESULTADOS	85
5.1	Estrategias de Asistencia.....	86
5.2	Estrategias de Prevención	87
5.2.1	<i>Ramp Metering</i>	88
5.2.2	Límites de Velocidad Variable (<i>Variable Speed Limits o VSL</i>).....	89
5.2.3	Paneles de Mensaje Variable y Mensajes Directos al Usuario	90
5.2.4	Estrategias de Diseño	91
6	CONCLUSIONES Y RECOMENDACIONES.....	93
6.1	Conclusiones Generales	93
6.2	Relevancia de los Estados de Tráfico.....	94
6.3	Relevancia de otras Variables de Control	96
6.4	Recomendaciones y Trabajo Futuro.....	97
	REFERENCIAS	100
	ANEXOS.....	108
	Anexo A: Otros Modelos de Predicción.....	109
	A-1 Principal Component Analysis (PCA) y Nonlinear Canonical Correlation Analysis (NLCCA)	109
	A-2 Redes Neuronales.....	109
	Anexo B: Modelos Ordinales	112
	Anexo C: Análisis de Correlación	114
	Anexo D: Cálculo de V_f	115
	Anexo E: Resultados de MNL 1:1 – 1:4.....	117
	E-1 Modelo Binario	117
	E-2 Modelo 3 DP	119
	E-3 Modelo 3 DT	123
	E-4 Modelo DP + DT.....	127
	Anexo F: Errores Estándar Estimados	131
	F-1 Modelo Binario	131
	F-2 Modelo 3 DP	131

F-3	Modelo 3 DT	132
F-4	Modelo DP + DT.....	132

ÍNDICE DE TABLAS

	Pág.
Tabla 2-1: Estudios Previos según Metodología y Tema	7
Tabla 2-2: <i>Crash Severity Studies</i> según Escala y Modelo	12
Tabla 3-1: Distribución de Accidentes	26
Tabla 3-2: Descripción de Variables de Espiras.....	28
Tabla 3-3: Variables de Información de Accidentes.	36
Tabla 3-4: Variables de Información Geométrica.	37
Tabla 3-5: Razón Caso-Control en la Literatura.....	39
Tabla 3-6: Bases de datos según Razón Caso-Control.	39
Tabla 4-1: Funciones de Propensión de la Alternativa “Accidente” – Modelo Binario.....	43
Tabla 4-2: Funciones de Propensión – Modelo 3 Alternativas DP.	43
Tabla 4-3: Funciones de Propensión – Modelo 3 Alternativas DT.	44
Tabla 4-4: Funciones de Propensión – Modelo 5 Alternativas.	45
Tabla 4-5: Resultados Generales de Estimación – ER (%).	48
Tabla 4-6: Resultados Generales de Estimación – FPR (%).	48
Tabla 4-7: Resultados de Estimación con 85% de muestra – Modelo Binario 1:5. ...	49
Tabla 4-8: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:5.	53
Tabla 4-9: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:5.	58
Tabla 4-10: Resultados de Estimación con 85% de muestra – Modelo DP + DT 1:5.....	62
Tabla 4-11: Riesgos Relativos Modelo 3 DP vs Modelo DP + DT.....	67
Tabla 4-12: Riesgos Relativos Modelo 3 DT vs Modelo DP + DT	69
Tabla 4-13: Resultados Validación – FPR y ER (%).	70
Tabla 4-14: Probabilidades y Ponderadores.	73
Tabla 4-15: Resultados de Estimación con muestra completa – MNLMp.....	77
Tabla 4-16: Resultados de Estimación con muestra completa – MLM Heterocedástico.....	78

Tabla 4-17: Resultados de Estimación con muestra completa – MLM Homocedástico.....	79
Tabla 8-1: Vf por día de semana.	116
Tabla 8-2: Resultados de Estimación con 85% de muestra – Modelo Binario 1:1.	117
Tabla 8-3: Resultados de Estimación con 85% de muestra – Modelo Binario 1:2.	117
Tabla 8-4: Resultados de Estimación con 85% de muestra – Modelo Binario 1:3.	118
Tabla 8-5: Resultados de Estimación con 85% de muestra – Modelo Binario 1:4.	118
Tabla 8-6: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:1.	119
Tabla 8-7: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:2.	120
Tabla 8-8: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:3.	121
Tabla 8-9: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:4.	122
Tabla 8-10: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:1.	123
Tabla 8-11: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:2.	124
Tabla 8-12: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:3.	125
Tabla 8-13: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:4.	126
Tabla 8-14: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:1.....	127
Tabla 8-15: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:2.....	128
Tabla 8-16: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:3.....	129
Tabla 8-17: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:4.....	130
Tabla 8-18: Errores Estándar Estimados con 85% de la muestra– Modelo Binario.	131
Tabla 8-19: Errores Estándar Estimados con 85% de la muestra – Modelo 3 DP.	131
Tabla 8-20: Errores Estándar Estimados con 85% de la muestra – Modelo 3 DT.	132
Tabla 8-21: Errores Estándar Estimados con 85% de la muestra – Modelo 5 Alternativas (DP+DT).....	133

ÍNDICE DE FIGURAS

	Pág.
Figura 2-1: Estructuras de un MNL y HLM.	18
Figura 3-1: Esquema de la Autopista Central.....	23
Figura 3-2: Esquema de nomenclatura para observaciones.....	24
Figura 3-3: Configuración de PM.....	27
Figura 3-4: Ejemplo para velocidad promedio ponderada por flujo.....	29
Figura 3-5: Ubicación Estación Meteorológica “Las Terrazas Oficina DGA”.	30
Figura 3-6: Diagrama Fundamental.....	32
Figura 3-7: Velocidades de Aguas Arriba (V1) vs Velocidades de Aguas Abajo (V2).	34
Figura 3-8: Proporción de Varianza Explicada en función del número de Clusters. .	35
Figura 3-9: Estados de Tráfico.	35
Figura 4-1: Evolución de FPR – Validación.	71
Figura 5-1: Cambios en los Estados de Tráfico.....	88
Figura 8-1: Esquema de una neurona artificial.....	110
Figura 8-2: Esquema del Modelo de Medida.	112
Figura 8-3: Gráfico de Velocidades medidas en PM007110-C1 el día 14/05/2012.	115

RESUMEN

El año 2012 se reportaron alrededor 19.500 accidentes en las zonas urbanas de la Región Metropolitana, Chile, de los cuales el 16% corresponde a accidentes en autopistas urbanas. Estos accidentes viales ocurren en distintas condiciones o Estados de Tráfico. Por condición o Estado de Tráfico se entiende una combinación única entre flujo, densidad y velocidad que determinan la operación de una vía. La finalidad de esta investigación es determinar la influencia que los Estados de Tráfico en autopistas tienen sobre la ocurrencia y severidad de los accidentes. Por severidad se consideran dos dimensiones: la severidad según daños personales (DP) y la disrupción en el tráfico (DT), usando la cantidad de pistas bloqueadas como *proxy* para DT.

Para este estudio se usó información de accidentes viales, flujos y velocidades instantáneas de circulación en Autopista Central, ubicada en Santiago de Chile, más un conjunto de variables de control relacionadas con la geometría y el clima. Por medio del algoritmo *k-means Clustering*, se definieron cuatro Estados de Tráfico usando las velocidades recolectadas por espiras magnéticas en la autopista, como una cantidad óptima.

Parte de la metodología de trabajo considera la calibración de un modelo de elección discreta *logit* a partir de datos reales, donde la ocurrencia (o no) de accidentes constituyen las alternativas, utilizando un muestreo de tipo Caso-Control. Los Estados de Tráficos se mostraron ser significativos, afectando la probabilidad de ocurrencia y severidad de los accidentes. Se concluye que los Estados de Tráfico de transición de velocidades desde altas a bajas o viceversa, son los que muestran un mayor riesgo en comparación con estados donde la velocidad se mantiene constante siendo alta o baja.

Con estos resultados, es posible la formulación de estrategias de mejor asistencia o prevención de accidentes que los operadores pueden considerar evaluar para su posterior implementación.

Palabras Clave: Accidentes, Estados de Tráfico, Modelos de Elección Discreta, Logit, Caso-Control, Autopistas Urbanas, Disrupción del Tráfico, Daños Personales.

ABSTRACT

The year 2012 around 19,500 accidents were reported in urban areas of the Metropolitan Region, Chile, of which 16% are urban highway accidents. These road accidents occur in different conditions or Traffic States. By condition or Traffic State means a unique combination of flow, density and speed that determine the operation of a road. The purpose of this research is to determine the influence that Traffic States in urban highways have on the occurrence and severity of accidents. By severity are considered two dimensions according to personal injury (PI) and traffic disruption (TD), using the road blockage as proxy for TD.

For this study, information on road accidents, flows and instantaneous speeds in Autopista Central, located in Santiago de Chile, plus a set of control variables related to the geometry and the weather was used. Through the k-means clustering algorithm, four Traffic States using the speeds collected by magnetic loops on the highway were defined as an optimum amount.

Part of the work methodology considers the calibration of a discrete choice logit model based on real data, where the occurrence (or not) of accidents are the alternatives, using a case-control sampling type. The Traffic States were statistically significant, affecting the likelihood of occurrence and severity of accidents. It is concluded that the Traffic States of speeds transition from high to low or vice versa, are the ones showing a greater risk compared to states where the speed remains constant being high or low.

With these results, the development of strategies to better assistance or prevention of accidents that operators may consider testing for later deployment is possible.

Keywords: Accidents, Traffic States, Discrete Choice Models, Logit, Case-Control, Urban Highways, Traffic Disruption, Road Blockage, Personal Injury.

1 INTRODUCCIÓN

En este capítulo se presenta, en primer lugar, una pequeña motivación de esta investigación, mencionando datos gruesos de la importancia que ha ido tomando la seguridad vial en Chile en el marco de lo que significa la prevención de la ocurrencia de accidentes de tránsito. Dicha importancia está relacionada específicamente, en el contexto de esta tesis, con la prevención de accidentes automovilísticos en autopistas urbanas, teniendo en cuenta mediciones históricas de: variables de tráfico, geometría de la vía y condiciones climáticas. Por lo mismo, esta investigación está principalmente centrada en la estimación de probabilidades para ciertos tipos de accidentes basada en Estados de Tráfico. En segundo lugar, se mencionan y explican los objetivos y alcances de esta investigación, para continuar finalmente, con la estructura del contenido de esta tesis.

1.1 Motivación

Actualmente, Chile presenta un amplio interés respecto al tema de seguridad vial, sumado al hecho que en la última década se ha notado una clara tendencia al crecimiento del número total de accidentes de tránsito ocurridos a lo largo del país¹. Si bien existen oscilaciones en la totalidad de accidentes desde 1972, entre los años 2006 y 2012 el crecimiento de la siniestralidad va en aumento, llegando hasta un total de 61.791 siniestros de tránsito en el 2012, a diferencia de los 44.839 recabados el 2006. Se podría pensar que el crecimiento del parque automotriz de Chile explica el crecimiento de siniestros, pero la tasa a la que los accidentes crecen es superior a la de los vehículos motorizados.

El año 2012 se reportaron 22.253 accidentes en las zonas urbanas de la Región Metropolitana, Chile, de los cuales (según una aproximación simple) el 14,2% corresponden a accidentes en autopistas urbanas (Américo Vespucio Norte, 2012;

¹ Datos obtenidos de la página web: http://www.conaset.cl/wp-content/uploads/2014/01/Evolucion_siniestros_transito_Chile1972_2012_1.xls

Américo Vespucio Sur, 2012; Autopista Central, 2012; Costanera Norte, 2012)². Este tipo de accidente es importante de ser estudiado, pues varios ocurren en un mismo lugar, o en condiciones similares de operación. Es decir, hay situaciones (o estados) de tráfico en las cuales la ocurrencia de un accidente se vuelve más probable. Si se logra encontrar la relación entre la ocurrencia de accidente y el estado de tráfico actual en la vía, sería posible entregar información, tanto a los usuarios como operadores, a fin de prevenir o asistir de manera más eficiente la ocurrencia de accidentes.

En estudios anteriores se han encontrado medidas probabilísticas de ocurrencias de accidente sin discriminar entre la severidad o tipo de accidente, mostrando una clara relación entre variables de tráfico, geometría de la vía y condiciones climáticas, con respecto a la ocurrencia (o no) de accidentes. Hay algunos estudios que sí consideran las condiciones de tráfico como variables que expliquen la ocurrencia de accidentes, pero tampoco discriminan de acuerdo a la severidad de cada accidente. Otros prefieren el uso de esta información para predecir o encontrar lugares muy probables de accidentes o *hotspots*.

Por otro lado, estudios que sí consideran diferentes tipos de accidentes clasificados según severidad estiman probabilidades con modelos de elección discreta. Sin embargo, no incluyen la alternativa que un accidente no ocurra, es decir, estiman probabilidades de que un accidente de cierto tipo ocurra dado que ya ha ocurrido un accidente, utilizando (algunas veces) variables que son inherentes a los conductores, lo cual es muy difícil de controlar por un operador de autopista concesionada, o bien, conseguir este tipo de información para efectos de usar el modelo. Además, consideran la severidad de un accidente solo según daños personales y no según la interrupción en el tráfico que provoca.

² Se registraron un total 3.152 accidentes en autopistas urbanas, donde las autopistas Américo Vespucio Norte Express, Américo Vespucio Sur, Costanera Norte y Autopista Central aportaron con 409, 449, 349 y 1942 accidentes respectivamente.

1.2 Objetivos e Hipótesis

La hipótesis de esta investigación es que los estados de tráfico son significativos en cuanto a la explicación de ocurrencia de accidentes de distinto tipo en autopistas.

Así, el objetivo general de esta investigación es formular un modelo que permita estimar, basado en las condiciones de tráfico observadas y otros factores fácilmente observables y no inherentes a los conductores, probabilidades de ocurrencia de accidentes. Los accidentes se diferencian en función de su severidad según daños personales y interrupción en el tráfico. Se desea considerar, además, la alternativa que no ocurra un accidente.

Entre los factores fácilmente observables y no inherentes a los conductores se incluyen: el clima y la geometría de la infraestructura vial.

Específicamente, se espera que condiciones de tráfico, donde se tengan velocidades altas y constantes a lo largo de un tramo específico, presenten el menor riesgo de ocurrencia de cualquier tipo accidente en comparación con otras condiciones de tráfico. Por ejemplo, si la condición de tráfico considerase velocidades bajas y constantes, se espera que la probabilidad de ocurrencia de un accidente poco severo, aumente. Por el contrario, se espera que la probabilidad de ocurrencia de accidentes con algún grado de severidad, disminuya.

Otros factores relacionados con el tráfico, se espera que sean estadísticamente significativos, afectando positivamente el valor de la probabilidad de ocurrencia de accidentes. Lo mismo se asume que ocurrirá en condiciones adversas de geometría y clima. *A priori*, no se sabe específicamente cómo pueden afectar en la probabilidad de ocurrencia de cada tipo de accidente.

1.3 Alcances

A continuación se presentan los alcances de esta investigación:

- El modelo se formula para los accidentes que ocurren en autopistas urbanas. En particular, se considera como vía en estudio la Autopista Central ubicada en Santiago de Chile, calibrando un modelo usando datos de ella.

- Los tipos de accidentes que el modelo incluye son: choques, colisiones y volcamientos, ya que estos son los que, según la intuición, pueden ser explicados por las condiciones de tráfico. No se consideran atropellos, por ejemplo.
- Se toman en cuenta solo aquellos accidentes ocurridos en la vía expresa, es decir, no aquellos que ocurren en caleteras, pistas de entrada y salida de la autopista.

1.4 Contenido

En el capítulo 2 se presenta la base teórica para la metodología empleada, además de la más común y frecuentemente utilizada en estudios revisados. Luego, se presenta el estado del arte en lo que respecta a estudios de seguridad vial relacionados con accidentes en autopistas. En el capítulo 3 se presenta en detalle el diseño experimental, explicando la metodología y el manejo de los datos para crear las muestras que se utilizaron para estimar los modelos. Finalmente en los capítulos 4 y 5 se presentan los resultados y conclusiones, respectivamente.

2 MARCO TEÓRICO

En la literatura se encuentran modelos de predicción de accidentes en tiempo real, que consideran la influencia de factores contribuyentes tales como las condiciones climáticas, la geometría de la vía y las condiciones del tráfico; en su ocurrencia. En lo que respecta a condiciones de tráfico, éstas se obtienen a partir de detectores ubicados en la vía.

Estos modelos, como cualquiera, presentan limitaciones en la predicción de ocurrencia de accidentes en tiempo real. Hossain & Muromachi (2012) exponen las siguientes:

- Ubicación del detector: el rendimiento de los modelos propuestos depende ampliamente de la ubicación de los detectores de datos de tráfico escogidos para un choque dado.
- Espacio de las variables: el potencial espacio de las variables de los estudios existentes, ha sido sustancialmente grande y diverso teniendo en cuenta que el tamaño de la muestra de accidente suele ser considerablemente pequeño.
- Métodos de modelación: los típicos modelos empleados para predicción de choques en tiempo real puede ser clasificados en: modelos estadísticos y de inteligencia artificial. Los primeros incluyen a *Matched Case-Control Logistic Regression*, donde se dejan de lado variables de tráfico por estar muy correlacionadas entre ellas. Los de inteligencia artificial, que utilizan redes neuronales, pueden utilizar variables correlacionadas como predictores, pero se esperan conocimientos previos suficientes (que no siempre se tienen para el periodo de modelación) con respecto al dominio expuesto del problema a través de la interrelación entre los predictores, para suplir la demanda de recursos que las redes neuronales requieren

Además de los modelos de predicción de accidentes, existen los modelos que permiten estudiar la severidad de los accidentes, calculando probabilidades de ocurrencia de ciertos tipos de accidentes dado que un accidente ya ha ocurrido. Es decir,

no predicen ocurrencia de accidentes. Para estudiar este ámbito se utilizan modelos de elección discreta.

A continuación, se expone la revisión bibliográfica que dicta el estado del arte sobre este tema de investigación, exponiendo aspectos importantes sobre estudios que utilizan estos modelos.

2.1 Revisión Bibliográfica

La literatura que se ha revisado y que está más relacionada con el tema de ocurrencia de accidentes en autopistas, se puede dividir en tres grandes temas. En primer lugar existen los estudios de **predicción de ocurrencia de accidentes**, que se han enfocado en la predicción de tasas de accidentes, o bien, en la estimación de algunas medidas probabilísticas para cuantificar el riesgo de la ocurrencia de accidentes en autopistas. En segundo lugar, otros estudios buscan definir o **predecir posibles hotspots** donde ocurran accidentes en las autopistas, utilizando varios conjuntos de variables de distinta índole. Finalmente, existen los **estudios de severidad** que estiman probabilidades de accidentes de distinto tipo, donde cada tipo se refiere a un nivel de severidad según lesión previamente definido, sin considerar la posibilidad de que un accidente no ocurra. Es decir, estiman probabilidades condicionales de ocurrencia de un cierto tipo de accidente dado que ya ha ocurrido alguno. En la Tabla 2-1, se presenta un listado resumen de los estudios previos organizados por tema y metodología empleada.

A continuación se describe lo reportado en estudios anteriores para establecer el estado del arte, justificar y respaldar los objetivos y alcances de esta investigación.

2.1.1 Estudios de Predicción de Ocurrencia de Accidentes

Este tipo de investigaciones corresponde a aquellos estudios que se han enfocado en la predicción de tasas de accidentes. Es decir, estimar algunas medidas probabilísticas en función de ciertas variables para cuantificar el riesgo de la ocurrencia de accidentes en autopistas.

Tabla 2-1: Estudios Previos según Metodología y Tema

Tema	Metodologías	Estudios Previos
Predicción de Ocurrencia	MCCLR	Abdel-Aty et al. (2004) Park & Oh (2009) Zheng et al. (2010) Zheng (2012) Xu et al. (2012) Ahmed et al. (2012a, 2012b)
		Golob & Recker (2003) Golob & Recker (2004) Golob et al. (2004a, 2004b)
	Redes Neuronales	Kononov, Durso, et al. (2012) Kononov, Hersey, et al. (2012)
	Otras Metodologías	Lee et al. (2002) Martin (2002) Oh et al. (2005) Wang et al. (2009) Yeo et al. (2013)
Identificación de <i>Hotspots</i>		Kononov (2002) Ivan (2004) Abdel-Aty & Pande (2004, 2005) Chung et al. (2009) Hossain & Muromachi (2012)
		Khattak et al. (2002) Kockelman & Kweon (2002) Abdel-Aty (2003) Abdel-Aty & Abdelwahab (2004) Yamamoto & Shankar (2004) Khorashadi et al. (2005) Eluru & Bhat (2007) Malyshkina & Mannering (2008) Pande & Abdel-Aty (2009) Haleem & Abdel-Aty (2010) Jung et al. (2010) Quddus et al. (2010) Anastasopoulos & Mannering (2011) Christoforou et al. (2011) Savolainen et al. (2011) Yang et al. (2011) Bham et al. (2012) Christoforou et al. (2012) Ye & Lord (2013)
Estudios de Severidad	Modelos de Elección Discreta	

Estas variables corresponden a características que se observan previamente a la ocurrencia de un accidente y se definen como los Precursores de Accidentes (Lee et al.,

2002). Éstos pueden ser: la variación de la velocidad, densidad vehicular, geometría de la vía y condiciones climáticas. Oh et al. (2005) también hace mención a estos precursores categorizándolos en cuatro grandes factores: ambiente, condiciones de tráfico, vehículos y conductores; y concluye que la detección de ellos cuando se tornan inestables e indeseables permiten la identificación anticipada de accidentes. Esta inestabilidad, específicamente, altas variaciones de la velocidad es considerado por Lave (1985) como el principal factor que conduce a accidentes.

En la literatura se menciona que los modelos de predicción de choques en tiempo real podrían aumentar su precisión si el modelo usa información de las características del tráfico tanto para situaciones donde sí hubo un accidente como para cuando no lo hubo (Caso-Control), mientras se controlan otros factores no inherentes a los conductores, como el clima y la geometría de la vía (Abdel-Aty et al., 2004). Otra manera de mejorar el rendimiento de los modelos es considerando la metodología de Caso-Control para ciertos tipos de accidentes y utilizando actualización bayesiana, como en el estudio de Ahmed et al. (2012a), donde se compara el rendimiento de predicción de accidentes entre un modelo que considera todos los tipos de choques y otro que solo considera choques por alcance. En Ahmed et al. (2011) se menciona que estos modelos se desempeñan mucho mejor en condiciones de congestión en comparación a condiciones de flujo libre.

Wang et al. (2009) investigan los efectos de la congestión de tráfico en accidentes por medio de la estadística utilizando distribuciones espaciales y no-espaciales Poisson-Lognormal y Poisson-Gamma, mientras se controlan otros factores contribuyentes como el flujo horario, largo de segmento de vía, número de pistas, radio de curvatura y gradiente.

Por otro lado, los modelos de predicción en tiempo real que dependen solo de información de tráfico, son útiles en autopistas con características geométricas poco riesgosas (Ahmed et al., 2012b), ya que una misma turbulencia en el tráfico, puede afectar al conductor de diferente manera en una autopista con geometrías adversas y bajo condiciones climáticas distintas.

Estos modelos retornan medidas probabilísticas que pueden ser utilizadas en el diseño de sistemas de información de advertencias, o bien, para evaluar la efectividad de varias estrategias de operación en términos de seguridad vial. Sin embargo, en Park & Oh (2009) se concluye que los modelos deben ser capaces de reaccionar con sensibilidad a las condiciones de tráfico, ya que los accidentes ocurren durante un periodo muy corto de tiempo debido a diversos elementos tanto directos como indirectos.

En otros estudios (Golob et al., 2004a, 2004b), que también estiman medidas probabilísticas, se desarrolla la herramienta FITS (*Flow Impacts on Traffic Safety*), que se puede utilizar para evaluar los cambios en tendencias de seguridad vial que resultan de cambios en el tráfico. El único *input* que FITS necesita son observaciones de intervalos de 30 segundos de mediciones provenientes de espiras individuales.

En Kononov, Durso, et al. (2012) se mencionan las *Safety Performance Functions* (SPF) diciendo que son la mejor medida de seguridad descrita por Hauer y Persaud (1997). Las SPF tratan de reflejar la compleja relación que existe entre la exposición, medida por el flujo anual promedio y conteos de accidentes, en una unidad de sección de vía en un periodo de tiempo. Por lo mismo, se calibran SPF para un corredor específico que relacionan tasas de accidentes con flujo-densidad horario y velocidad obtenidos en estaciones de registro automático, proponiendo un algoritmo para un sistema de límite de velocidad variable (*Variable Speed Limits* o VSL), que se puede utilizar como sistema de gestión de seguridad vial en autopistas.

Finalmente, Yeo et al. (2013) calculan dos tipos de probabilidades de ocurrencia de accidentes para Estados de Tráfico. Según vehículo-milla-recorrida y vehículo-hora-recorrida, considerando datos de espiras obtenidos durante tres años agregados en intervalos de cinco minutos.

2.1.2 Estudios de Identificación de *Hotspots*

La identificación de *hotspots* se refiere a encontrar lugares donde sea muy probable que ocurran accidentes. Identificar estos lugares puede ayudar a encontrar falencias en las vías para mejorar el diseño de la vía o señalización.

Mediante un diagnóstico directo y reconocimiento de patrones, Kononov (2002) estipula que uno de los principales usos de las SPF es identificar lugares que experimenten más accidentes que los esperados, demostrando un potencial para reducción de los mismos.

En otro estudio se menciona que la inclusión de la medida vehículo-milla-recorrida por año en un análisis de choques en autopistas es un importante paso para la identificación de lugares riesgosos (Ivan, 2004), concluyendo que la incidencia de un choque no varía linealmente con el flujo, al ser éste representado como un conteo de exposición (número de ensayos), intensidad del tráfico y niveles de servicio (flujo sobre capacidad).

Abdel-Aty & Pande (2004 y 2005) buscan *hotspots*, dadas las condiciones de tráfico que preceden un accidente. Para esto utilizan *Probabilistic Neuronal Networks* (PNN) con variables de entrada seleccionadas a partir de un análisis básico de Caso-Control, utilizando razones de riesgo.

De manera similar, Hossain & Muromachi (2012) investigan la forma de identificar lugares riesgosos mediante los modelos *Random Multinomial Logit* (RMNL) y *Bayesian Belief Net* (BBN). Ellos, además de considerar solo el tráfico en los detectores, introducen (para capturar la variación espacial) una nueva variable con una nomenclatura del tipo “dabX”, donde a y b representan los detectores (aguas arriba y aguas abajo), y X la variable en consideración, por ejemplo, el flujo, la velocidad, la densidad u ocupación. También, se introduce el índice de congestión como la diferencia entre la velocidad a flujo libre y la velocidad del detector, sobre la velocidad a flujo libre.

Finalmente, para dar paso a un modelo más complejo, Chung et al. (2009) proponen un método continuo de monitoreo de autopistas para detectar lugares con altas concentraciones de choques, mediante un enfoque de perfil de riesgo continuo (CRP). A diferencia de otros modelos, CRP no requiere de segmentación de la autopista, y la correlación espacial de las colisiones no afecta los resultados.

2.1.3 Estudios de Severidad de Accidentes

En la literatura se ha estudiado la severidad de accidentes de tránsito, pero muchas investigaciones no toman en cuenta, dentro de sus alternativas de accidente según severidad, una que considere la no ocurrencia de accidentes. Es decir, se estiman probabilidades de ocurrencia de un accidente con un nivel de severidad especificado previamente, dado que un accidente ya ha ocurrido.

Para obtener estas probabilidades se pueden utilizar dos clases de modelos de elección discreta de tipo Nominal u Ordinal, según la escala de la variable dependiente. Dentro de cada una de estas clases existen más tipos de modelos que han sido utilizados en estudios pasados (Tabla 2-2).

El objetivo principal de la mayoría de los estudios revisados es identificar los factores que contribuyen a la severidad de la lesión sufrida en accidentes. Estos factores ingresan al modelo como las variables explicativas y se estiman los parámetros asociados a éstas. Existen cuatro tipos de factores, que son comúnmente utilizados, tales como: datos específicos de cada accidente (modelo y año del vehículo, sexo y edad del conductor, ingesta de alcohol, entre otros), condiciones climáticas, estado y geometría de la vía, y condiciones del tráfico.

Las investigaciones previas utilizan, en la estimación de sus modelos, información de todos los factores mencionados (Anastasopoulos & Mannering, 2011; Khorashadi et al., 2005; Malyshkina & Mannering, 2008), pero en algunos casos no es así, sino que solo una parte es usada (Abdel-Aty & Abdelwahab, 2004; Christoforou et al., 2011, 2012; Eluru & Bhat, 2007; Khattak et al., 2002; Quddus et al., 2010). Por ejemplo, Kockelman & Kweon (2002) solo consideran datos específicos de cada accidente, mientras que otros (dependiendo de sus objetivos) van agregando no solo otros factores contribuyentes como el clima, tráfico, y geometría y estado de la vía, sino que también, más características propias del accidente, tales como: uso del cinturón de seguridad del conductor (Abdel-Aty, 2003), lesión del conductor y pasajero más perjudicado (Yamamoto & Shankar, 2004), distancia de frenado y deficiencia de seguimiento al momento del accidente (Jung et al., 2010).

Tabla 2-2: *Crash Severity Studies* según Escala y Modelo

Escala	Modelo	Estudios Previos
Nominal	MNLM	Khorashadi et al. (2005) Malyshkina & Mannering (2008) Pande & Abdel-Aty (2009) Anastasopoulos & Mannering (2011) Bham et al. (2012)
		Abdel-Aty & Abdelwahab (2004) Haleem & Abdel-Aty (2010)
	Probit	Haleem & Abdel-Aty (2010) Christoforou et al. (2011) Christoforou et al. (2012)
Ordinal	OPM	Khattak et al. (2002) Kockelman & Kweon (2002) Abdel-Aty (2003) Haleem & Abdel-Aty (2010) Yang et al. (2011)
	OLM	Abdel-Aty (2003) Quddus et al. (2010)
	Otros	Yamamoto & Shankar (2004)* Eluru & Bhat (2007)** Jung et al. (2010)*** Quddus et al. (2010)****

*Modelo Probit Ordinal Bivariado (BOPM)

**Modelo Logit conjunto con coeficientes correlacionados aleatorios para una respuesta binario-ordenada. Podría llamarse, Modelo Logit Conjunto Aleatorio Ordinal Bivariado.

***Modelos policotómicos de regresión logística ordinal y secuencial.

****OLM parcialmente restringido o PC-GOLOGIT.

En Savolainen et al. (2011) se presenta un resumen sobre estudios de este tema y critican el hecho que la severidad según lesión se discretiza en niveles³ tales como: lesión fatal, discapacidad, sin discapacidad, posible lesión y sin lesión. Estos niveles, en el marco de los modelos de elección discreta, deben seguir una escala ordinal al estar naturalmente ordenados, lo que indica qué tipo de modelos son apropiados para representar este fenómeno (Abdel-Aty, 2003; Jung et al., 2010; Kockelman & Kweon,

³ Estos niveles, en algunos estudios, se rigen por la escala KABCO: *fatal injury or killed* (K), *incapacitating injury* (A), *non-incapacitating injury* (B), *possible injury* (C), y *property damage only* (O).

2002). Khorashadi et al. (2005) mencionan las desventajas de usar un MNLM por sobre un HLM, pues se dejaría de lado la posible correlación entre alternativas al usar estos niveles de severidad, en particular. Malyshkina & Mannering (2008), por otro lado, critican el uso de modelos ordinales, pues los modelos con escala nominal relajan una restricción impuesta por los primeros de que una variable no puede crecer simultáneamente para distintas categorías de severidad.

La estimación de parámetros y correlación espacial y temporal son otros asuntos que mencionan Savolainen et al. (2011), pues en la mayoría de la literatura revisada se estiman parámetros fijos, y estos restringen los efectos de las variables explicativas a ser iguales a lo largo de todas las observaciones individuales de severidad. Esto no es realista debido a la heterogeneidad que puede existir en la población de viajeros involucrados en accidentes, ya que existen diferencias en toma de riesgos, y factores psicológicos, entre otros. Esto puede ser solucionado al estimar modelos con parámetros aleatorios (Anastasopoulos & Mannering, 2011; Eluru & Bhat, 2007). Por otro lado, es diferente la situación cuando cada observación corresponde a un choque y no a cada individuo afectado por un accidente, utilizando para la estimación solo variables explicativas no inherentes a los individuos.

La presencia de correlación espacial y temporal se puede dar cuando dos o más accidentes ocurren cercanos en el tiempo y espacio, por lo que no tomar en cuenta esto para la estimación del modelo, se estarían dejando de lado factores comunes que no han sido observados, perdiendo eficiencia y precisión de estimación.

Eluru & Bhat (2007) estudiaron la severidad por lesión y el uso de cinturón de seguridad. En este trabajo mencionan la sustancial inflación que puede existir en el uso de cinturón de seguridad al ignorar el método de selección de muestra (*Sample Selection*). Este método se puede utilizar para agregar una alternativa a estos modelos que sea la correcta implementación de una alternativa que sea un no accidente.

Finalmente, Ye & Lord (2013) examinaron los efectos del tamaño muestral para tres modelos de severidad según lesión comúnmente estimados: MNLM, OPM y MLM; usando simulación de Montecarlo. Además, se mencionan los modelos nominales (MNLM, HLM y MLM) y ordinales (OLM, OPM y Modelo Logit Ordinal Mixto) más

comunes. Se recomienda que los tamaños muestrales mínimos deben ser 1000 observaciones para el OPM, 2000 para el MNLM y 5000 MLM.

A continuación, se describe brevemente el detalle teórico de las metodologías más empleadas.

2.2 Modelos de Predicción

Los modelos de predicción de ocurrencia de accidentes más comunes son: la Regresión Logística con Emparejamiento Caso-Control (*Matched Case-Control Logistic Regression* o MCCLR), Análisis de las Componentes Principales (*Principal Component Analysis* o PCA) junto a un Análisis de Correlación Canónica No Lineal (*Nonlinear Canonical Correlation Analysis* o NLCCA), y Redes Neuronales.

Para efectos de esta investigación se presenta la teoría del primer modelo junto a un análisis crítico que expone ventajas y limitaciones. Si el lector desea indagar más sobre el resto de los modelos, puede dirigirse al Anexo A: Otros Modelos de Predicción.

2.2.1 Matched Case-Control Logistic Regression (MCCLR)

La regresión logística corresponde a un análisis de regresión utilizado para predecir el resultado de una variable categórica en función de variables explicativas, que pueden ser tanto categóricas como continuas. El uso de este método sirve para estimar probabilidades de ocurrencia de un evento en función de otros factores. Estas probabilidades describen el posible resultado de la variable dependiente de un único ensayo en función de las variables explicativas, utilizando la función *logit*.

En el contexto de accidentes, la regresión logística busca determinar la probabilidad de accidente en función de una serie de variables explicativas. Para lograr esto, se deben definir casos y controles. Los casos corresponden a situaciones en donde sí ocurrió un accidente, y los controles son condiciones similares pero libres de éstos. Así, la variable dependiente es binaria, que posee dos posibles resultados: Accidente (casos) y no accidente (controles).

En la Ecuación (2-1) se muestra la forma funcional de este modelo, que representa la probabilidad de ocurrencia de un accidente.

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i})}} \quad (2-1)$$

β_0 es la constante independiente y β_j es el parámetro que corresponde al logaritmo de la Razón de Momios (*Odds Ratio*) y explican el efecto aditivo de una unidad de cambio en la j-ésima variable explicativa, cuando el resto de las variables permanecen fijas. Por su parte, la razón de momios también se conoce como Razón de Riesgo (*Hazard Ratio*) y se define como la razón en las probabilidades de que un accidente ocurra y no ocurra. Esta razón, se obtiene aplicando la función exponencial al parámetro.

La Ecuación (2-2) corresponde la forma lineal sobre la cual se realiza el análisis de regresión. Esta se obtiene aplicando logaritmo sobre la razón entre la Ecuación (2-1) y la unidad menos la misma ecuación:

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i} \quad (2-2)$$

Comentarios

El MCCLR es un modelo simple de estimar que demuestra tener buenos resultados, apuntando, precisamente, a calcular la influencia que tienen las variables explicativas sobre las probabilidades de ocurrencia de un accidente. Esta influencia, usando esta metodología, se puede interpretar a partir de las *Odds Ratio* estimadas por el modelo, o bien, por alguna otra medida en particular.

Además, las salidas del modelo corresponden a medidas probabilísticas de ocurrencia de accidentes, y pueden ser utilizadas en el diseño de sistemas de información de advertencias, o bien, para evaluar la efectividad de varias estrategias de operación en términos de seguridad vial.

Sin embargo, el modelo asume efectos lineales de las variables explicativas, lo cual no es necesariamente correcto. También, se asume que los parámetros son fijos cuando, en realidad, podrían tener una distribución de probabilidad asociada. Esto se podría solucionar utilizando actualización bayesiana, o bien, estimando los parámetros con algún otro modelo que permita aleatoriedad de los mismos.

Otra limitación de este modelo, es que se omiten variables por estar muy correlacionadas con otras. Por la relación fundamental, las variables de tráfico si lo están. Esto se soluciona utilizando las condiciones de tráfico.

2.3 Modelos de Elección Discreta

Como se mencionó, algunos estudios buscan estimar probabilidades de accidentes según severidad por lesión, dado que un accidente ya ha ocurrido. Para estimar las probabilidades, en la mayoría de los estudios se utiliza un enfoque econométrico, es decir, se utilizan modelos de elección discreta debido a que la variable de respuesta es de carácter discreto. Dentro de los modelos más utilizados, se tienen los que consideran una escala nominal para la variable dependiente y los que consideran una naturaleza ordinal para la misma.

Los modelos de escala nominal que más se utilizan son los modelos Logit Multinomial (MNL), Logit Jerárquico (HLM) y Logit Mixto (MLM). Y los de escala ordinal son: Logit Ordinal (OLM) y Probit Ordinal (OPM). Existen otros modelos que han sido empleados, pero que no son tan utilizados como los mencionados.

A diferencia de la Regresión Logística con Emparejamiento Caso-Control, los modelos de elección discreta permiten variables dependientes con más de dos categorías, y por lo tanto se pueden hacer diferencias entre tipos de accidentes previamente definidos, como se hace en esta tesis.

Este tipo de modelos, usualmente están sustentados bajo la teoría de la utilidad aleatoria, pero en este caso se toma un enfoque distinto. La función de utilidad deja de llamarse de esta manera para dar paso a la función de propensión a la severidad. En los siguientes párrafos se explica de manera detallada la adaptación de este enfoque.

La Teoría de la Utilidad Aleatoria afirma que la utilidad se deriva de los atributos de los bienes y no de los bienes *per se*. En este caso, se afirma que la propensión se deriva de los precursores de un accidente y no del accidente propiamente tal. Si se considera un conjunto $\underline{A}$ de alternativas (tipos de accidente) y un conjunto $\underline{X}$ de precursores, un accidente q perteneciente a un conjunto Q (de accidentes totales en la muestra) tiene un conjunto de precursores $\underline{x} \in \underline{X}$ y un subconjunto de alternativas

$\underline{A}(q) \in \underline{A}$. Cada alternativa A_i tiene asociada una propensión T_{iq} para el accidente q , que es la suma de dos componentes:

$$T_{iq} = V_{iq} + \varepsilon_{iq} \quad (2-3)$$

En la expresión anterior, V_{iq} es la componente representativa, en función de los precursores $\underline{x}$, y viene dada por:

$$V_{iq} = \sum_{k=1}^K \theta_{ik} x_{ik} \quad (2-4)$$

Por otra parte, ε_{iq} corresponde la componente aleatoria del modelo, que refleja las particularidades de cada accidente, los errores de medición y observación del modelador.

En esta investigación se trabajará con modelos que consideran una escala nominal para la variable de respuesta, y a continuación se describe la teoría de cada uno de ellos. La teoría de los modelos *Logit* y *Probit* ordinal, comunes en la literatura, se presenta en el Anexo B: Modelos Ordinales.

2.3.1 Modelo Logit Multinomial (MNL)

El modelo Logit Multinomial se deriva a partir de la suposición de que los errores (ε) descritos en la teoría de la utilidad aleatoria (propensión), distribuyen con funciones IID Gumbel, llegando al siguiente del modelo (McFadden, 1974):

$$P_i = \frac{\exp(\beta V_i)}{\sum_{A_j \in \underline{A}(q)} \exp(\beta V_j)} \quad (2-5)$$

P_i es la probabilidad de “escoger” el bien i . Una de las principales propiedades de este modelo es la independencia de las alternativas irrelevantes:

$$P_i/P_j = \exp[\beta(V_i - V_j)] \quad (2-6)$$

Esta propiedad muestra que es posible conocer la proporción de elección entre dos alternativas, sin conocer las utilidades de las alternativas restantes. Sin embargo, esto es una desventaja cuando hay dos alternativas que tienen correlación en los errores. Para tratar esta limitación existe el HLM.

2.3.2 Modelo Logit Jerárquico (HLM)

De la Figura 2-1 (Ortúzar, 2000) se aprecia que el modelo Logit Jerárquico, a diferencia del MNL, presenta estructuras que consideran alternativas correlacionadas dentro de un mismo nido, pero no permite correlación entre alternativas pertenecientes a nidos diferentes.

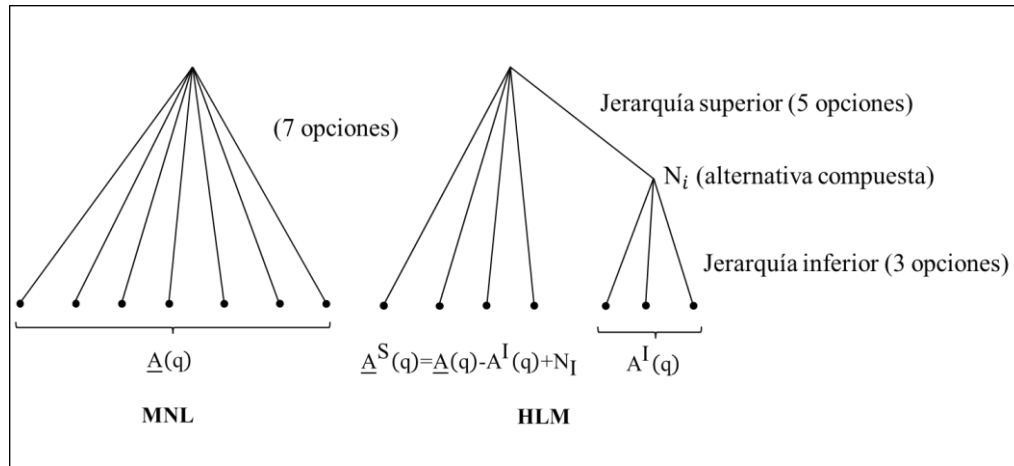


Figura 2-1: Estructuras de un MNL y HLM.

La utilidad del nido N_i posee dos componentes: la EMU, que corresponde a la utilidad máxima esperada (o bien, propensión) entre alternativas del nido inferior (Ecuación (2-7)), y el vector W de los precursores comunes entre las alternativas del nido (V_i' representa a los precursores diferentes).

$$EMU = \ln \sum_{A_j \in A^I(q)} \exp(V_i') \quad (2-7)$$

De esta forma, la utilidad del nido N_i puede expresarse como:

$$\tilde{U}_i = \phi \cdot EMU + \alpha \cdot \underline{W} \quad (2-8)$$

Para que el modelo sea consistente, ϕ debe cumplir que:

$$0 < \phi_1 \leq \phi_2 \leq \dots \leq \phi_s \leq 1 \quad (2-9)$$

La Ecuación (2-9) indica que ϕ_i debe disminuir a medida que el nivel de jerarquía i decrece.

En el caso sencillo de dos niveles jerárquicos, sin atributos comunes para las alternativas de la jerarquía inferior $A^I(q)$, la probabilidad que un accidente q corresponda a un accidente del tipo $A_i \in A^I(q)$ viene dada por la Ecuación:

$$P_{iq} = \frac{\exp(\tilde{U}_q)}{\sum_{j \in (A^S(q) - N_I)} \exp(\phi V_{jq}) + \tilde{U}_q} \cdot \frac{\exp(V_{iq})}{\sum_{j \in A^I(q)} \exp(V_{jq})}, \quad (2-10)$$

donde $A^S(q)$ corresponde las alternativas de la jerarquía superior y $\tilde{U}$ la utilidad de la alternativa compuesta N_I . Por otro lado, la probabilidad de elegir $A_i \in A^S(q)$ es:

$$P_{iq} = \frac{\exp(\phi V_{iq})}{[\sum_{j \in (A^S(q) - N_I)} \exp(\phi V_{jq})] + \tilde{U}_q} \quad (2-11)$$

El problema que presenta este modelo, es que no permite correlación entre alternativas que se encuentran en nidos distintos. Para solucionar esta limitación, se puede usar el MLM.

2.3.3 Modelo Logit Mixto (MLM)

El modelo Logit Mixto se considera superior al MNLM y HLM, puesto que es capaz de agregar correlación entre alternativas, tanto del mismo nido como en distintos. Además, permite agregar heterogeneidad en las preferencias individuales, algo que para efectos de esta investigación no nos concierne.

En el caso de agregar correlación entre alternativas, se incluyen componentes de error en las funciones de propensión (utilidad). Es decir, las funciones, además de poseer una parte determinística y un error con distribución Gumbel, pueden ser modificadas agregando términos de error con una cierta distribución especificada.

Por ejemplo, en el caso de un modelo que considera cinco alternativas y lo siguientes pares de alternativas representan la correlación entre ellas: (2-3), (4-5), (2-4) y (3-5). La formulación del modelo debería ser como la expresada en la Ecuación (2-12).

$$\begin{aligned} T_1 &= V_1 + \mu_j + \eta_0 + \varepsilon_1 \\ T_2 &= V_2 + \mu_I + \eta_1 + \varepsilon_2 \\ T_3 &= V_3 + \mu_I + \eta_2 + \varepsilon_3 \\ T_4 &= V_4 + \mu_H + \eta_1 + \varepsilon_4 \\ T_5 &= V_5 + \mu_H + \eta_2 + \varepsilon_5 \end{aligned} \quad (2-12)$$

Así, este modelo lo que hace es agregar más componentes aleatorias a las funciones de propensión del MNL, para representar el efecto de correlación entre alternativas.

La estimación del MLM es mucho más compleja en comparación con los MNLM. En general, se utiliza máxima verosimilitud simulada (Ortúzar y Willumsen 2011), como se hará en esta investigación.

2.4 Indicadores y Tests Estadísticos

En esta sección se describen indicadores y pruebas estadísticas, que ayudan a cuantificar la interpretación de los parámetros, el desempeño de los modelos y la comparación entre ellos.

2.4.1 Riesgo Relativo (*Relative Risk*)

El riesgo relativo (RR) corresponde a la razón entre la probabilidad de un evento ocurriendo cuando un factor se encuentra presente, y la probabilidad del mismo evento ocurriendo cuando ese factor no está. En este caso los factores están representados por variables binarias.

Para el caso de variables continuas, el RR se interpreta cuando la variable aumenta en una unidad.

En esta investigación, se obtiene un riesgo relativo promedio, calculando para cada observación i el riesgo relativo y obteniendo el RR promedio con todas las observaciones.

Así, la manera de calcular el riesgo relativo promedio con respecto a la probabilidad de ocurrencia de un accidente del tipo j , se expresa en la Ecuación (2-13).

$$\overline{RR}_j = \sum_i \frac{P_{ji}^1}{P_{ji}^0} \quad (2-13)$$

P_{ji}^1 y P_{ji}^0 representan las probabilidades de ocurrencia del accidente tipo j en la observación i , con y sin el factor presente, respectivamente.

2.4.2 Porcentaje Correctamente Predicho o FPR (*First Preference Recovery*)

El FPR es a una medida agregada que corresponde a las veces en que el modelo predice correctamente el accidente ocurrido, es decir, es la proporción de accidentes que efectivamente resultan en la opción de máxima propensión de acuerdo al modelo.

No obstante, según Ortuzar (2000), un modelo con muy alto valor de FPR debiera rechazarse igual que uno con muy bajo FPR, y para entender esto se deben definir: el valor esperado de FPR (ER), como el promedio de las probabilidades de ocurrencia de los tipos de accidente para cada observación individual; y el valor esperado de la proporción al azar CR (*Chance Recovery*), como el promedio de las probabilidades de ocurrencia al azar de cada alternativa para todas las observaciones. Como todas las observaciones tienen la misma cantidad de tipos de accidente disponibles a ocurrir, CR es igual a E(CR).

Dado lo anterior, se puede comparar el FPR, para un modelo en específico, con el ER y E(CR). Si FPR y ER son similares y mayores que E(CR) se puede decir que el modelo es razonable e informativo.

2.4.3 Test de Razón de Verosimilitud

En Ortuzar (2000) se describe este *test*. Para llevar a cabo esta prueba se debe calcular el estadístico LR (Ecuación (2-14)) que distribuye χ^2 con r grados de libertad (o el número de restricciones lineales), y la comparación entre modelos se da cuando uno es la versión restringida del otro.

$$LR = -2\{l(\hat{\theta}_r) - l(\hat{\theta})\} \quad (2-14)$$

La hipótesis nula se rechaza si $LR > \chi^2_{r;95\%}$, lo que indica que el modelo restringido es erróneo.

3 DISEÑO EXPERIMENTAL

La metodología de trabajo considera la calibración de modelos de elección discreta a partir de datos reales, donde la ocurrencia (o no) de accidentes constituyen alternativas como variable dependiente. Se desea investigar la evolución de la complejidad de especificación del modelo. Así, se estima desde un modelo *Logit* Binario hasta uno *Logit* Multinomial de cinco alternativas, donde siempre se considera una alternativa que corresponde a la no ocurrencia de un accidente. Estas alternativas, relativas a la ocurrencia de un accidente, se definen en función de la severidad, en donde ésta no está relacionada solo con daños personales (DP), sino que también con el tipo de disrupción que provoca en el tráfico (DT). Para este último se utiliza la cantidad de pistas que quedaron bloqueadas luego de ocurrido el accidente, como *proxy* de disrupción.

Por último, se estiman modelos *Logit* Mixto para tratar las limitaciones que los modelos *Logit* Multinomial poseen para modelar este fenómeno, y verificar si una estructura de componentes de error representa la posible correlación existente entre alternativas de accidente.

Como toda calibración (o estimación) de modelos es necesario tener una base de datos, donde cada observación corresponda a un accidente (o bien, no accidente) con sus características observadas previas a la hora de ocurrencia (condiciones de tráfico, geometría de infraestructura, condiciones climáticas).

En el resto de este capítulo se describe la construcción de la base de datos utilizada para calibrar los modelos antes mencionados. Primero, se describe brevemente el proyecto de infraestructura Autopista Central. Luego, se define la nomenclatura para construir cada observación de la base de datos. Finalmente, se presenta un análisis de los datos que corresponde a explicar en qué consiste cada uno de los tipos de información disponibles para después crear la base de datos que se utilizó en la fase de estimación.

La especificación, estimación y resultados de los modelos se presentan en el Capítulo 4.

3.1 Descripción General de Autopista Central

El proyecto de infraestructura vial Autopista Central se encuentra ubicado en Santiago de Chile. La autopista atraviesa la capital chilena en sentido Norte-Sur, por lo tanto se clasifica como una autopista urbana.

Autopista Central posee dos ejes: General Velásquez con una extensión cercana a los 21 kilómetros y Ruta 5 con una extensión de 39,5 kilómetros aproximadamente (Figura 3-1). También, consta con un sistema de telepeaje en flujo libre, mediante un dispositivo electrónico asociado a cada vehículo por su placa patente, que permite el cobro de peajes sin necesidad de que los vehículos se detengan.

En la gran mayoría de su extensión, Autopista Central posee, en promedio, tres pistas de circulación por sentido, salvo en algunos tramos del eje General Velásquez donde la cantidad de pistas disminuye a dos.

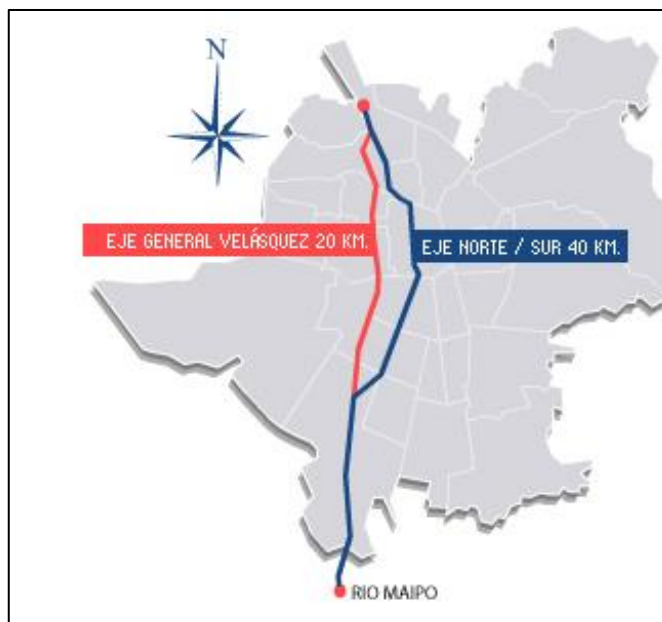


Figura 3-1: Esquema de la Autopista Central.

Además, la infraestructura de Autopista Central posee un equipamiento muy avanzado destinado, por ejemplo, a dar asistencia e información a los usuarios, y a medir información de tráfico en tiempo real con el uso de espiras magnéticas.

3.2 Definición de observaciones

Antes de describir la información y la confección de la base de datos, se define la nomenclatura que constituye la forma en la que se construyó cada observación. Es importante destacar que, a diferencia de los estudios de demanda o econométricos, cada observación no corresponde a un individuo sino un accidente con características definidas por la información recabada. Estas características corresponden a factores no inherentes a los conductores (como lo son las condiciones climáticas, el tráfico y la geometría de infraestructura vial) en la ocurrencia de accidentes.

Para atribuirle características a cada accidente se utilizó el esquema presentado en la Figura 3-2. En ella se especifican tres estaciones. La primera es el grupo de detectores que constituyen la Espira 1 ubicada aguas arriba del lugar del accidente, la cual corresponde a la Estación 2. La Estación 3 corresponde a la Espira 2 aguas abajo del accidente en sentido de avance de los vehículos.

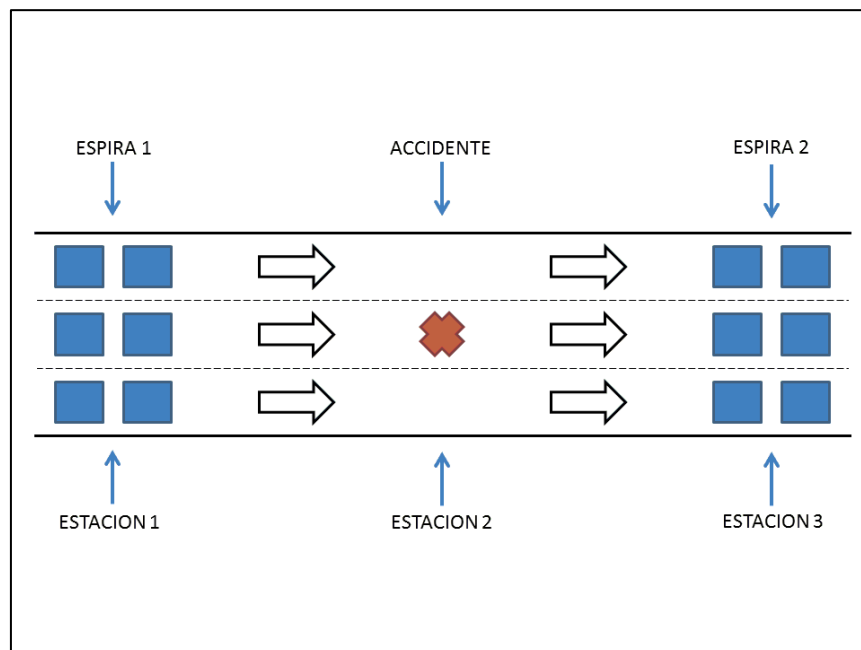


Figura 3-2: Esquema de nomenclatura para observaciones.

Como se mencionó, esta investigación considera dentro de las alternativas de cada modelo (*choice set*), la posibilidad que no ocurra un accidente. Por esto se adoptó la

definición de Controles utilizada en estudios que se basan en la investigación epidemiológica de Caso-Control (explicada en Schlesselman (1982)). Por lo tanto, Casos (accidentes) como Controles (no accidentes) se rigen bajo la misma nomenclatura explicada.

3.3 Base de Datos

La base de datos a utilizar considera cuatro tipos de información: de accidentes, de tráfico, del clima y de geometría de la vía. A continuación, se detalla cada una de ellas, para luego describir las variables explicativas, Estados de Tráfico y diseño Caso-Control utilizado.

3.3.1 Accidentes

Comparado con reportes de Carabineros de Chile, se espera que los datos provistos por Autopista Central sean más completos y presenten menor error por registros mal recopilados. Errores que, en estudios previos (Pande & Abdel-Aty, 2009), han demostrado ser importantes. Esto se debe a que Autopista Central posee un sistema de información muy eficiente que registra cualquier tipo de incidente que pueda entorpecer potencialmente el flujo normal de los vehículos en la vía, y además posee sistemas de asistencia a los usuarios.

Se cuenta con el registro de 1.978 accidentes ocurridos en Autopista Central durante el año 2012 y la primera semana de Enero 2013. Para cada uno de los accidentes la información con que se cuenta incluye:

- Hora, fecha de ocurrencia y duración.
- Tipo de evento: Atropello, Caída, Choque, Colisión, Otro tipo de accidente y Volcamiento.
- Ubicación del accidente, especificando Eje, Vía (Local o Expresa), Calzada y Punto Kilométrico.
- Distribución de daños personales, es decir, cantidad de personas ilesas, heridas y víctimas fatales en el lugar.
- Distribución de pistas bloqueadas, considerando pistas de vías expresa, pista de enlace, berma y pistas de entrada o salida de la autopista.

Como no todos los eventos registrados se encuentran dentro de los alcances de la investigación, se utilizaron sólo aquellos que corresponden a choques, colisiones y volcamientos ocurridos en la vía expresa y que bloquean pistas de la misma, considerando que si se bloquea la berma, no existe disrupción en el tráfico. Con esto se obtiene un total de 1.219 accidentes o casos.

Definición de Tipos de Accidente

A partir de la información de accidentes se definieron los tipos de accidentes que se utilizaron en los modelos de elección discreta calibrados, considerando dos dimensiones de severidad. La severidad dada por daños personales (DP) y por disrupción en el tráfico (DT). En la literatura, lo más común, es discriminar los niveles de severidad solo en función de DP.

Para daños personales se definen dos niveles: accidentes sin heridos ni víctimas fatales (A_I) y accidentes con al menos un herido o muerto (A_H). La disrupción en el tráfico también considera dos niveles: un bloqueo menor de pistas (una o ninguna, A_1p) y un bloqueo mayor (dos o más pistas, A_2p).

En la Tabla 3-1 se muestran las distribuciones de accidentes para cada dimensión de severidad por separado y la conjunta de DP y DT.

Tabla 3-1: Distribución de Accidentes

		DP		TOTAL
		A_I	A_HM	
DT	A_1p	49,8%	25,3%	75,1%
	A_2p	11,5%	13,4%	24,9%
TOTAL		61,3%	38,7%	100,0%

3.3.2 Información de Espiras

Autopista Central posee, a lo largo de ambos de sus ejes, Puntos de Medida (PM) o Espiras que agrupan Detectores. La cantidad de Detectores por PM depende de la cantidad de pistas, es decir, si existen, por ejemplo, tres pistas entonces PM tiene tres

Detectores. El equipamiento de la autopista, en cuanto a Espiras, consta de 349 Puntos de Medida y 716 Detectores.

Como se mencionó en la sub-sección anterior, se utilizan aquellos accidentes que hayan ocurrido en la vía expresa, por lo tanto no sirven aquellos PM ubicados en las pistas de entra o salida de la autopista, quedando un total de 196 Espiras con datos a utilizar.

Cabe destacar que todos los Puntos de Medida poseen una configuración doble (Figura 3-3). Ésta implica una conexión en serie de dos dispositivos que constituyen un detector, y esto permite minimizar el error de medición de la velocidad, ya que la distancia entre dispositivos es fija y solo se necesitan los instantes de tiempo en que un vehículo acciona los dispositivos.

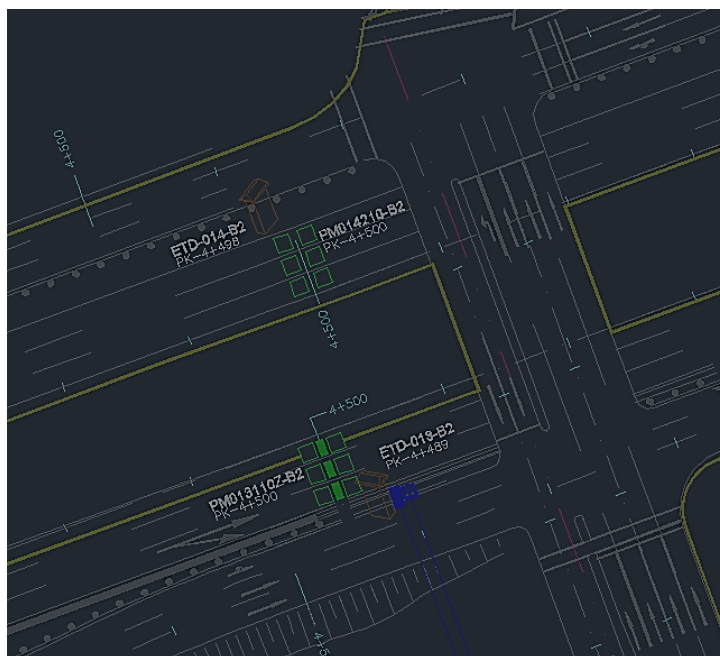


Figura 3-3: Configuración de PM.

En la Tabla 3-2 se detallan los campos presentes en la base de datos correspondiente a la medición de las espiras. La información de velocidad, ocupación y flujo se agrega cada un minuto. La velocidad tiene como unidades de medida kilómetros por hora. La ocupación es la cantidad de segundos dentro del minuto de agregación en

que el detector estuvo ocupado. El flujo corresponde a los vehículos por hora que se registran cada minuto, es decir, si el flujo es de 60 vehículos por hora, el conteo fue de un vehículo en ese minuto.

Tabla 3-2: Descripción de Variables de Espiras

Campo	Descripción
PM	Código de Punto de Medida
DETECTOR	Código de Detector
PK	Punto kilométrico para cada PM
NOM_SENTIDO	Corresponde a la calzada donde se ubica PM: Oriente o Poniente
FECHA	Fecha y hora de la medición
VELOCIDAD	Velocidad promedio de todos los vehículos que pasan por un detector cada minuto [km/hr]
OCUPACION	Tiempo en que el detector está ocupado con respecto al intervalo de medición [seg]
FLUJO	Flujo medido en un minuto [veh/hr]

Los datos corresponden a las mediciones del año 2012 exceptuando el mes de Septiembre. En total se tienen 343.780.141 mediciones. Es necesario mencionar que los detectores no funcionan correctamente todo el tiempo.

Para cumplir con los fines de esta investigación, es necesario agregar la información medida por los Detectores asociados a un Punto de Medida como si fuesen un solo Detector. Para esto, sólo son necesarios el flujo y la velocidad.

En el caso del flujo, éste solo se debe sumar, pero en el caso de la velocidad se definen: medias, desviaciones estándar y coeficientes de variación. Dentro de las medias calculadas de la velocidad, se calculó el promedio ponderado por el flujo, puesto que no es sensato pensar en un promedio simple, cuando los detectores pueden reportar mediciones muy distantes entre ellas, es decir, en una pista se puede tener un flujo alto y en otra uno despreciable.

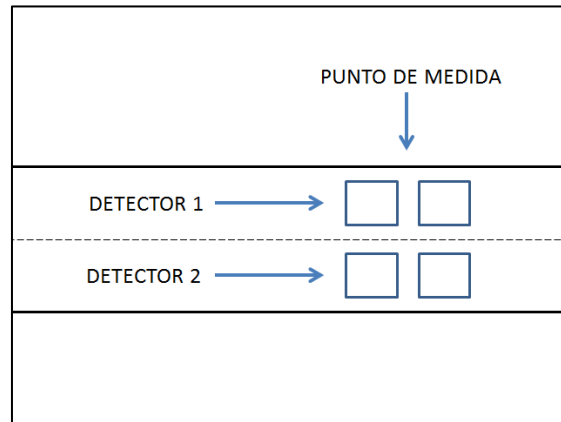


Figura 3-4: Ejemplo para velocidad promedio ponderada por flujo.

Por ejemplo, si se tiene el caso simple de dos pistas (Figura 3-4), donde en un minuto se midió los pares velocidad-flujo: 60-120 y 90-240. Y sabiendo que cada detector mide una velocidad promedio de las velocidades de cada vehículo que circuló por esa pista, es decir, $\frac{\sum v_{i,j}}{N_j}$ que corresponde a la suma de todas las velocidades de los vehículos que circularon por la pista j , dividido por la cantidad de vehículos. Y, el flujo (Q_j) como a la cantidad de vehículos N_j multiplicado por 60. El cálculo del promedio ponderado de la velocidad para dos pistas, sería de la siguiente manera:

$$\frac{\frac{\sum v_{i,1}}{N_1} \cdot Q_1 + \frac{\sum v_{i,2}}{N_2} \cdot Q_2}{Q_1 + Q_2} = \frac{60 \cdot 120 + 90 \cdot 240}{120 + 240} = 80$$

que es distinto de 75, que corresponde al promedio simple entre 60 y 90.

3.3.3 Información de Geometría e Infraestructura

La geometría e infraestructura de la vía en estudio se obtuvo a partir de una inspección visual realizada sobre planos CAD del proyecto Autopista Central y la aplicación *Google Street View*TM.

Para recopilar la información se definieron tramos, que corresponde al segmento de vía entre un par de Espiras consecutivas espacialmente. Por lo tanto, los valores que toman las variables corresponden a características físicas de los tramos o presencia de elementos, *a priori*, distractores para los conductores. Por ejemplo, se definen variables

como CURVA, que indica la presencia de curvas en un tramo. La totalidad de las variables definidas se presentan en la Sección 3.3.5.

En la autopista existen 193 tramos que cumplen con la nomenclatura para la definición de observaciones, pero en solo 162 de ellos ocurrió al menos un accidente durante el 2012.

3.3.4 Información de Condiciones Climáticas

La información climática se obtuvo a partir de mediciones recopiladas en la estación meteorológica “Las Terrazas Oficinas DGA”, a cargo de la Dirección General de Aguas (DGA) del Ministerio de Obras Públicas (MOP). Esta estación se encuentra ubicada en la zona central de Santiago, cercana al lugar donde Autopista Central cruza con el Eje Libertador Bernardo O’Higgins (ver Figura 3-5). Se utilizó solo una estación pues no existían estaciones cercanas a la autopista por el sector norte y las estaciones cercanas por el sector sur no reportaban la información necesaria o no poseían las mediciones para el año de estudio.

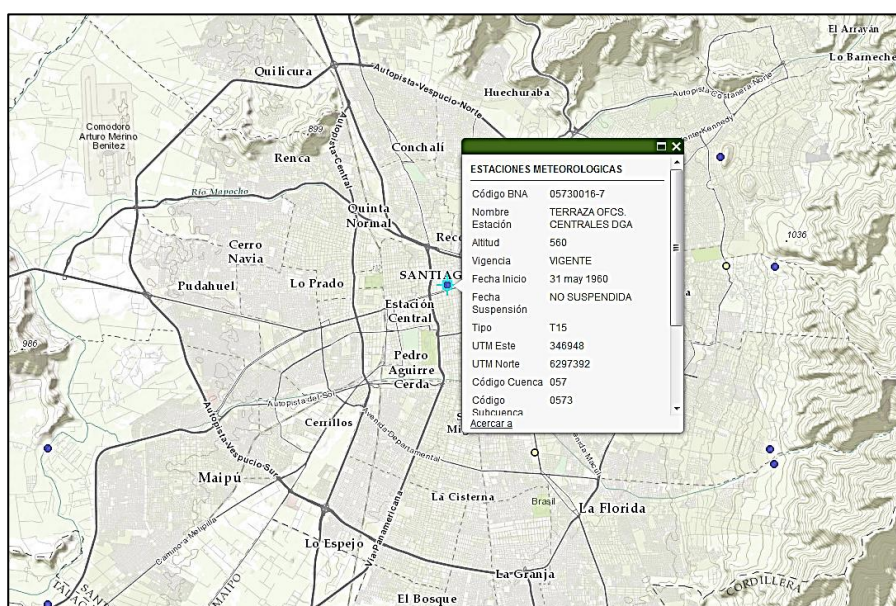


Figura 3-5: Ubicación Estación Meteorológica “Las Terrazas Oficina DGA”.

Esta información disponible corresponde a mediciones de humedad relativa, temperatura y precipitación, con una resolución de datos horaria para todo el año 2012.

Se asumió que los datos obtenidos en esta estación eran constantes a lo largo de la autopista. Para verificar este supuesto se realizó un análisis de correlación entre estaciones que se encuentran al poniente y oriente de la estación de Las Terrazas (ver Anexo C: Análisis de Correlación).

3.3.5 Estados de Tráfico y Variables Explicativas

En esta sub-sección se definen los Estados de Tráfico y las variables explicativas presentes en las bases de datos finales, a partir de los tipos de información presentados anteriormente.

i. *Estados de Tráfico*

Los Estados de Tráfico corresponden a condiciones del tráfico en una vía, y teóricamente se definen como una combinación única entre las tres variables macroscópicas de tráfico: flujo (q), velocidad (v) y densidad (k). La definición de cada una de estas variables corresponde a:

- El flujo (q) corresponde a la cantidad de vehículos que cruzan un punto de la vía por unidad de tiempo.
- La velocidad (v), específicamente, la velocidad media espacial es la media aritmética de las velocidades instantáneas de los vehículos que se encuentran en la vía en un momento dado.
- La densidad (k) corresponde a la cantidad de vehículos por unidad de distancia dentro de un tramo de la vía.

Las tres variables macroscópicas cumplen una relación algebraica denominada Relación Fundamental, donde el flujo (q) corresponde al producto entre la velocidad media espacial (v) y la densidad (k). Lighthill & Whitham (1955) fueron los primeros en dar indicios de esta relación, proponiendo que el flujo es una función de la densidad. Por otro lado, Greenshields (1935) fue el primero en verificar empíricamente la Relación Fundamental encontrando la primera aproximación gráfica a lo que corresponde el Diagrama Fundamental q - k , donde cada punto (k, q) de la curva representa un Estado de Tráfico y la pendiente de la recta que une ese punto con el origen corresponde a la velocidad de los vehículos en ese estado (Figura 3-6).

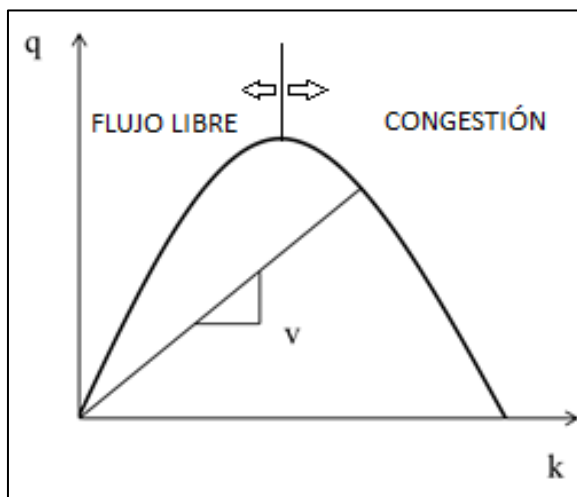


Figura 3-6: Diagrama Fundamental.

Se busca definir Estados de Tráfico para cada tramo que represente la condición de tráfico en los cinco minutos previos a la hora del accidente.

Dado que cada Estado de Tráfico presenta una velocidad distinta, se utilizará esta variable para caracterizarlo. Así, la variable velocidad se discretiza en grupos o *clusters* que representan los Estados, utilizando el mismo algoritmo *K-means Clustering* que se utiliza en otros estudios. Esto se hace así, pues como se mencionó anteriormente, se espera que la variable velocidad medida por las espiras presente un bajo error debido a la configuración de espiras dobles instaladas en la Autopista. Otra razón para utilizar solo la velocidad, es la posible correlación, existente entre las macro-variables.

En la literatura revisada se presentan varios métodos para definir los Estados de Tráfico. Por ejemplo, en el *Highway Capacity Manual* (TRB, 2010) se dice que la operación en autopistas se divide en seis Estados de Tráfico basados en la densidad, asignando niveles de servicio (LOS) a cada uno. Wu (2002), por otro lado, expone la existencia de cuatro estados de acuerdo al comportamiento del seguimiento vehicular y la posición relativa de los vehículos: tráfico fluido libre, tráfico fluido agrupado, tráfico congestionado agrupado y tráfico congestionado detenido.

En el ámbito de estudios previos sobre seguridad vial, tenemos las siguientes definiciones de Estados de Tráfico:

- Golob & Recker (2004) definen 8 “regímenes” de tráfico: *Light free flow*, *heavily congested flow*, *congested flow*, *light right-lane-variable flow*, *flow at capacity*, *heavy variable flow*, *heavy steady flow*, and *flow near capacity*.
- Zheng (2012) utiliza la *Wavelet Transformation* (WT) para agrupar las condiciones de tráfico en tres estados: Flujo Libre, Transición y Congestión.
- Xu et al. (2012) considera la variable densidad para encontrar cinco grupos que representan los estados, mediante el algoritmo *K-means Clustering*.
- Yeo et al. (2013) definen, mediante inspección visual, cuatro Estados de Tráfico basados en las velocidades aguas arriba y aguas abajo del incidente: *Free Flow* (FF), *Bottleneck* (BN), *Back of Queue* (BQ) y *Congestion* (CT).

Antes de aplicar el algoritmo *K-means Clustering* para encontrar los Estados de Tráfico, es necesario calcular la Velocidad de Flujo Libre (V_f), pues existen minutos de medición por cada detector en la autopista iguales a cero, para las variables: velocidad, ocupación y flujo. Esto significa que en ese minuto, no pasó nadie, lo que, en teoría, no corresponde a una velocidad igual a cero, sino que corresponde a V_f . Por lo tanto, es necesario calcular V_f de la Autopista Central.

El detalle del cálculo de V_f se encuentra en el Anexo D: Cálculo de V_f , y se llegó a un valor de 93 km/hr.

Teniendo todo esto en consideración, se procede en implementación del algoritmo *K-means Clustering* para encontrar los estados de tráfico presentes en la autopista de interés. Se decide utilizar este método pues es simple y mejor que una inspección visual.

K-means Clustering consiste en un método de agrupación de datos particionando una matriz de n observaciones en k grupos o *clusters*, en donde cada observación pertenece al *cluster* con la media o controides más cercano. Esta medida de cercanía se calcula usando la distancia euclidiana.

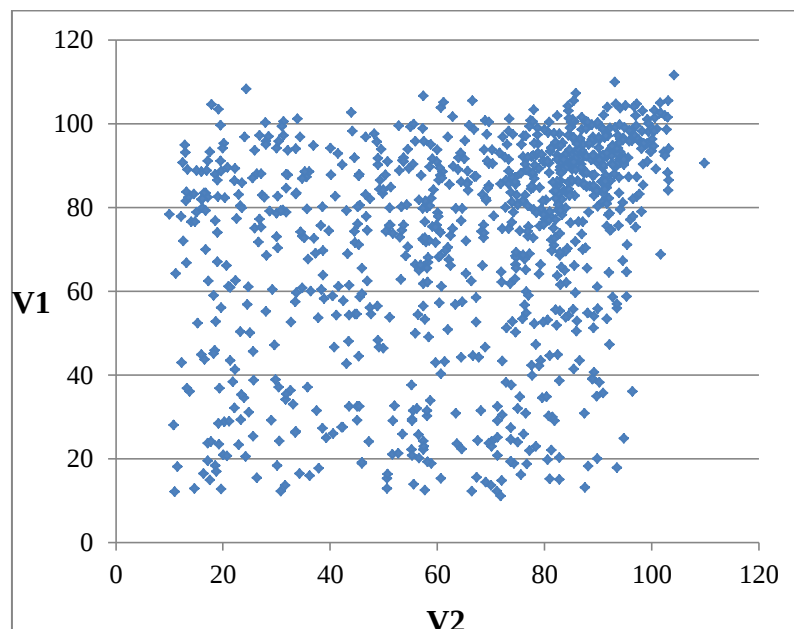


Figura 3-7: Velocidades de Aguas Arriba (V1) vs Velocidades de Aguas Abajo (V2).

Para aplicar este algoritmo y definir los Estados de Tráfico presentes tiempo antes de ocurrido cada accidente, se calculó, para una sub-muestra de 927 accidentes, una media ponderada por el flujo con las mediciones de velocidad y flujo de los cinco minutos anteriores a la hora registrada para cada accidente, considerando las estaciones de Aguas Arriba (V1) y Aguas Abajo (V2). Con esto se obtiene un vector de 927×2 , que se da como *input* al *software* MATLAB para particionar este vector en k grupos. En la Figura 3-7 se observa una mayor concentración de puntos cuando V1 y V2 fueron altas, lo que *a priori* indica un estado de Flujo Libre.

Al ser k un número arbitrario, éste se varió desde 1 hasta 10 y con el método del codo (*Elbow Method*) se definieron cuatro Estados de Tráfico como una cantidad óptima (Figura 3-8). Este método consiste en elegir un número de grupos tal que al agregar otro no entregue información importante para la modelación de los datos, es decir, se escoge la cantidad de *clusters* tal que al agregar otro la ganancia marginal de proporción de varianza explicada decrece. En la Figura 3-9 se exponen los cuatro Estados de Tráfico, donde el Estado 1 (de transición) O “Final de Cola” corresponde a velocidades altas para V1 y bajas para V2. El Estado 2 corresponde a uno de Flujo Libre con velocidades altas

tanto para V1 como para V2. El Estado 3 es un estado congestionado donde las velocidades son bajas para ambas V1 y V2. Y finalmente, el Estado 4 (de transición) o “Cuello de Botella” corresponde a uno donde V1 son velocidades bajas y V2 altas. También se muestra el centroide de cada *cluster*.

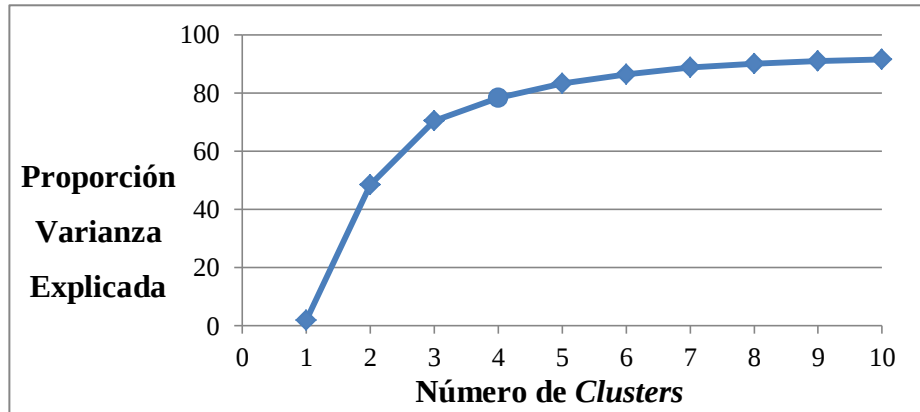


Figura 3-8: Proporción de Varianza Explicada en función del número de Clusters.

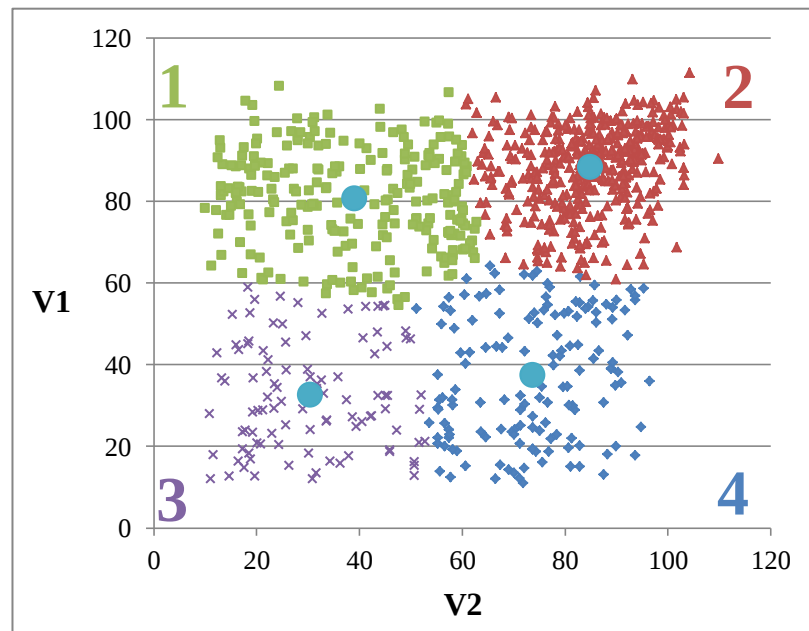


Figura 3-9: Estados de Tráfico.

ii. *Variables Explicativas*

En primer lugar, con la información de accidentes se definieron cinco variables del tipo *dummy* que se explican en la Tabla 3-3.

Tabla 3-3: Variables de Información de Accidentes.

Variable	Descripción
SEMANA	1 cuando el accidente ocurrió un día de semana y 0 en otro caso
DURANTE EL DÍA	1 cuando el accidente ocurrió durante el día y 0 cuando ocurrió de noche
ESTACIÓN	1 cuando el accidente ocurrió en Verano o Primavera y 0 en otro caso
HORA PUNTA	1 cuando el accidente ocurre en periodo punta mañana y 0 en otro caso. La punta mañana se define entre las 7 y 9 am
ACCIDENTE ACTIVO	1 si existe otro accidente que no haya sido despejado y haya ocurrido en el mismo lugar en la misma fecha y 0 en otro caso. Se define activo si ocurre dentro del intervalo de tiempo definido por la hora del accidente previo y su duración

En segundo lugar, se presenta en la Tabla 3-4, las variables obtenidas a partir de la información geométrica.

En tercer lugar, a partir de los datos climáticos, se definió la variable *dummy* LLUVIA, la cual toma el valor 1 cuando la estación meteorológica reportó algún valor de precipitación horaria, y 0 en otro caso.

Finalmente, en cuarto lugar se tienen las variables de tráfico definidas a partir de la información de las Espiras. Con esta información, fue posible extraer las mediciones de tráfico de las Estaciones 1 y 3 del diseño experimental desde $t=0$ hasta $t-30$ minutos, donde $t=0$ corresponde a la hora registrada para cada accidente. Estos 30 minutos de información fueron usados para definir 22 intervalos de tiempo, donde 11 corresponden a cada Estación. Los 11 intervalos se componen de 6 intervalos disjuntos de cinco minutos y 5 intervalos crecientes en el tiempo, es decir, desde $t=0$ hasta $t-10$, $t=0$ hasta $t-15$ y así sucesivamente hasta $t-30$.

Tabla 3-4: Variables de Información Geométrica.

Variable	Descripción
DISTANCIA ENTRE ESTACIONES 1 Y 3	Representa el largo del tramo en metros
PANEL MENSAJE VARIABLE	1 si el tramo presenta Panel de Mensaje Variable y 0 en otro caso
PASARELA	1 si el tramo presenta alguna Pasarela y 0 en otro caso
PASO SOBRE NIVEL	1 si el tramo presenta Paso Sobre Nivel y 0 en otro caso
PASO BAJO NIVEL	1 si el tramo presenta Paso Bajo Nivel y 0 en otro caso
PENDIENTE EN SUBIDA	1 si el tramo presenta una Pendiente en Subida y 0 en otro caso
PENDIENTE EN BAJADA	1 si el tramo presenta una Pendiente en Bajada y 0 en otro caso
PÓRTICO	1 si el tramo presenta Pórtico de cobro de peaje y 0 en otro caso
CRUZA PASO SOBRE NIVEL	1 si el tramo presenta una sección a nivel pero existe un paso sobre nivel que cruza el tramo y 0 en otro caso
CRUZA PASO BAJO NIVEL	1 si el tramo presenta una sección a nivel pero existe un paso bajo nivel que cruza el tramo y 0 en otro caso
DISMINUCIÓN CAPACIDAD	1 si el tramo presenta una disminución de capacidad por disminución de pistas y 0 en otro caso
RECTA	1 si el tramo es Recto sin curvas y 0 en otro caso
CURVA	1 si el tramo presenta Curvas y 0 en otro caso
SALIDA	1 si el tramo presenta una Salida de la vía expresa y 0 en otro caso
ENTRADA	1 si el tramo presenta una Entrada a la vía expresa y 0 en otro caso
MERGE	1 si el tramo corresponde a la convergencia de dos vías expresas y 0 en otro caso
TUNEL	1 si el tramo presenta túnel y 0 en otro caso

Para cada intervalo se calcularon cinco variables continuas de tráfico: la velocidad media (VM), la velocidad media ponderada por el flujo (VMP), la desviación estándar

de la velocidad (DE), el coeficiente de variación de la velocidad con media simple (CV) y el coeficiente de variación de la velocidad con la media ponderada (CVP).

Además, con VMP para todos los intervalos y Estaciones, se pudo obtener los Estados de Tráfico presentes cinco minutos antes de ocurrido cada accidente, como variables *dummy*.

3.3.6 Caso-Control

Esta investigación, como se mencionó, sigue la teoría de los estudios epidemiológicos de Caso-Control. Así, los accidentes son considerados como Casos y los no-accidentes como Controles.

Para la confección de la base de datos de los Casos, se utilizaron las variables descritas en la sub-sección anterior para caracterizar cada accidente. De un total de 1.219, 14 fueron eliminados por no cumplir con la nomenclatura de definición de observaciones, es decir, fueron accidentes ocurridos en los extremos de la autopista y no ocurrían dentro de un tramo (par de espiras espacialmente consecutivas pertenecientes a Autopista Central). Otros 278 accidentes, también fueron eliminados por carecer de información de Espiras (Septiembre 2012) o por tener información errónea en las mediciones. Con esto se obtiene un total de 927 Casos.

Por otro lado, tenemos la definición de los Controles, que están asociados a un Caso por el lugar físico de ocurrencia del accidente y se caracterizan con las mismas variables que los Casos. Los Controles son aquellas situaciones observadas una semana antes y después de cada Caso, considerando una holgura temporal de un rango de $t \pm 60$ minutos con respecto a la hora reportada para cada Caso. Con esta holgura es posible obtener situaciones en $t-60$, $t-30$, $t=0$, $t+30$ y $t+60$. Con esta definición obtenemos, *a priori*, 10 candidatos a ser Control para cada Caso, siempre comprobando que ningún Control coincida con la hora reportada para un Caso.

Al igual que como se hizo para los Casos, se debieron eliminar registros para los Controles debido a errores o carencia de información. Esto genera que no todos los Casos posean la misma cantidad de candidatos a ser Control.

Otro asunto es la razón numérica que debe existir entre la cantidad de Casos y Controles. Como se sabe, la cantidad de accidentes es fija, por lo tanto la razón Caso-Control está netamente definida por la cantidad de Controles que se considere por Caso, escogidos aleatoriamente. No existe respaldo teórico que muestre qué razón es la que se debe utilizar. Algunos autores recomiendan razones entre 1:3 hasta 1:5, por presentar errores estándar menores de los parámetros estimados. En la Tabla 3-5 se aprecia un cuadro resumen que muestra la razón Caso-Control y las cantidades utilizadas en algunos estudios previos.

Tabla 3-5: Razón Caso-Control en la Literatura.

Estudio	Razón	Casos	Controles
Abdel-Aty et al. (2004)	1:5	375	1875
Park & Oh (2009)	1:1	939	939
Zheng et al. (2010)	1:4	82	328
Ahmed et al. (2012a)	~1:3	30	880
Ahmed et al. (2012b)	1:4	447	1788
Xu et al. (2012)	1:4	278	1112

Tabla 3-6: Bases de datos según Razón Caso-Control.

Razón	Casos Útiles	Casos Inútiles	Total Observaciones
1:1	916	11	1832
1:2	914	13	2742
1:3	913	14	3652
1:4	912	15	4560
1:5	910	17	5460

La razón Caso-Control parece ser algo específico de cada investigación y debe ser probado. Por lo tanto, para esta tesis se confeccionaron cinco bases de datos variando la razón Caso-Control. Dado que la cantidad de candidatos a ser Control no es la misma para todos los Casos, el total de Casos utilizados tampoco fue siempre el mismo para las distintas bases de datos. Por ejemplo, si la razón fuese 1:1 esto indica que todos los 916 Casos útiles poseen al menos un candidato a ser Control. En cambio, si la razón fuese 1:2 todos esos Casos que poseían un solo candidato a ser Control pasan a ser inútiles. La Tabla 3-6 muestra los detalles para este aspecto.

4 ESTIMACION DE MODELOS Y RESULTADOS

Los modelos de elección discreta están sustentados en la teoría de la utilidad aleatoria. En nuestro caso se toma un enfoque distinto, en donde la función de utilidad se renombra como función de propensión a la severidad.

Como se menciona en el Capítulo 2, la teoría de la utilidad aleatoria afirma que la utilidad se deriva de los atributos de los bienes y no de los bienes *per sé*. En este caso, se afirma que la propensión se deriva de los precursores de un accidente y no del accidente propiamente tal.

Esta sección detalla la especificación y estimación de los modelos Logit Multinomial (MNL) y Logit Mixto (MLM), presentando, para cada uno de estos: las funciones de propensión, los signos esperados de los parámetros, y los resultados de la fase de estimación y validación.

Se destaca, que cada una de las bases de datos (según razón Caso-Control) fue dividida en dos partes. Una para la fase de estimación y otra para validar. Para obtener la parte que corresponde a estimación, se obtuvo aleatoriamente el 85% de los Casos y luego se asociaron sus Controles respectivos. El restante 15% corresponde a los Casos que quedaron fueran de la selección aleatoria y luego sus respectivos Controles fueron asociados. Por ejemplo, si tomamos la base de datos con una razón Caso-Control de 1:5. Ésta contiene un total de 5460 observaciones, de las cuales 910 corresponden a Casos y 4550 a Controles. Así, el 85% de 910 corresponde aproximadamente a 774 Casos, y asociando sus respectivos Controles tenemos una muestra para estimación de 4644 que cumple con la razón Caso-Control inicial. El restante 15% corresponde a la muestra de validación.

En primer lugar, se utilizó el MNL como modelo exploratorio, debido a su simpleza y rápida estimación, con el fin de estudiar la complejidad de especificación de los modelos según la definición de severidad, y también evaluar el desempeño de los modelos a medida que la razón Caso-Control cambia.

En segundo lugar, se utilizó el MLM para probar componentes aleatorias en las funciones de propensión que representen la correlación existente entre alternativas

simulando un *Cross-Nested Probit*, y además utilizar ponderadores para cada observación que corrijan la estimación de los parámetros, dado que el tipo de muestreo corresponde a uno del tipo “Caso-Control”.

Se intentó estimar modelos Logit Jerárquicos (HLM), pero la estimación arrojaba (en todo los casos) que el nido de jerarquía inferior colapsaba al de jerarquía superior.

La estimación de todos los tipos de modelos se realizó mediante el *software* BIOGEME (Bierlaire, 2003).

4.1 Logit Multinomial (MNL)

La estimación de modelos *logit* se llevó a cabo con el fin de investigar la evolución en la complejidad de especificación de los modelos, considerando las dos dimensiones de severidad definidas (daños personales y interrupción en el tráfico).

Esta complejidad de especificación se deriva en los siguientes cuatro tipos de modelos, donde siempre se considera como una de las alternativas, la posibilidad de que no ocurra un accidente:

- Modelo Binario, es decir, modelo que considera dos alternativas para la variable de respuesta. Accidente y No-Accidente.
- Modelo de 3 alternativas que define severidad solo con daños personales (DP).
- Modelo de 3 alternativas que define severidad solo con interrupción del tráfico (DT).
- Modelo de 5 alternativas que define severidad como DP + DT.

Un supuesto a destacar es que la función de propensión para la alternativa de no-accidente (Controles), se asume igual a cero. Ésta es la alternativa base.

Además, **se estimaron parámetros específicos de las variables para cada alternativa**. Esto se explica de la siguiente manera. Si, por ejemplo, se deseara estimar un modelo de elección de un bien cualquiera, la base de datos para estimar tendría, en cada observación i , una variable de respuesta y , por ejemplo, una variable de costo donde se presentarían los costos de todos los bienes a elección disponibles para el individuo i . Estimando un modelo *logit*, se podrían calcular las diferencias de los costos y estimar un solo parámetro de costo en común para todas las funciones de utilidad de

las alternativas. Sin embargo, en este caso eso no ocurre y por lo tanto se deben estimar parámetros específicos para cada alternativa.

Las funciones de propensión para los mejores modelos estimados, se presentan a continuación.

4.1.1 Funciones de Propensión

Las funciones de “utilidad”, o bien de propensión de cada alternativa en un modelo, pueden variar. Esto se debe a las distintas bases de datos utilizadas, definidas por la razón Caso-Control, utilizadas para la estimación. Esta variación se explica por el hecho de que la cantidad de Casos útiles fue variando en cada base de datos según razón Caso-Control, ya que, por ejemplo, en el caso de una razón de 1:1 se poseen más Casos que en una razón de 1:5 (Tabla 3-6). Es decir, se eliminan observaciones a medida que se aumenta la cantidad de Controles por Caso, lo que hace variar que parámetros estimados fueron significativos dependiendo de la base de datos utilizada.

Por otro lado, dado que el valor de la función de propensión para las alternativas “No-Accidente” es igual a cero, solo se muestran las funciones del resto de las alternativas en cada modelo, utilizando las mismas variables definidas en la Sección 3.3.5.

Otro punto importante es que en todos los modelos se utilizó, como Estado de Tráfico base, el Estado 2, que corresponde al estado de flujo libre.

El modelo binario tiene como opciones de variable dependiente, la ocurrencia de un accidente *versus* la no ocurrencia. En la Tabla 4-1 se muestran las funciones de propensión para este modelo.

El modelo con 3 alternativas definidas en función de la severidad solo según DP, tienen las siguientes alternativas: “Accidente sin heridos ni víctimas fatales” (A_I), “Accidente con al menos un herido o muerto” (A_HM) y “No-Accidente”. La Tabla 4-2 muestra sus funciones de propensión correspondientes.

Tabla 4-1: Funciones de Propensión de la Alternativa “Accidente” – Modelo Binario.

Razón	Variables en Función de Propensión
1:1 – 1:5	SEMANA DISTANCIA ENTRE ESTACIONES 1 Y 3 DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA PENDIENTE EN BAJADA x ESTADO FINAL DE COLA PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA LLUVIA x CURVA (*) CONSTANTE ACCIDENTE

(*) La variable “LLUVIA x CURVA” pertenece solo a la función con 1:1.

Tabla 4-2: Funciones de Propensión – Modelo 3 Alternativas DP.

Alt.	Razón	Variables en Función de Propensión
A_HM	1:1 – 1:5	ENTRADA DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA DURANTE EL DÍA EN DÍA DE SEMANA PENDIENTE EN BAJADA x ESTADO FINAL DE COLA LLUVIA x CURVA (*) MERGE (**) CONSTANTE A_HM
A_I	1:1 – 1:5	DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA LLUVIA x CURVA (***) DISTANCIA ENTRE ESTACIONES 1 Y 3 (****) CONSTANTE A_I

(*) La variable “LLUVIA x CURVA” pertenece solo a la función con 1:1 de la alternativa A_HM.

(**) La variable “MERGE” pertenece solo a las funciones con las razones desde 1:2 hasta 1:5 de la alternativa A_HM.

(***) La variable “LLUVIA x CURVA” pertenece solo a la función con 1:1 de la alternativa A_I.

(****) La variable “DISTANCIA ENTRE ESTACIONES 1 Y 3” pertenece solo a las funciones con las razones desde 1:2 hasta 1:5 de la alternativa A_I.

El otro modelo de 3 alternativas define la función de severidad solo según DT. Así, las alternativas son las siguientes: “Accidente con bloqueo menor de pistas” (hasta 1 pista bloqueada, A_1p), “Accidente con bloqueo mayor de pistas” (2 o más pistas bloqueadas, A_2p), y “No-Accidente”. Las funciones se ven en la

Tabla 4-3.

Tabla 4-3: Funciones de Propensión – Modelo 3 Alternativas DT.

Alt.	Razón	Variables en Función de Propensión
A_1p	1:1 – 1:5	DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA PENDIENTE EN BAJADA x ESTADO CONGESTIÓN PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA CONSTANTE A_1p
A_2p	1:1 – 1:5	SEMANA DISTANCIA ENTRE ESTACIONES 1 Y 3 PENDIENTE EN SUBIDA DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA LLUVIA x CURVA CONSTANTE A_2p

Finalmente, el modelo de 5 alternativas utiliza las dos dimensiones de severidad (DP y DT). Las alternativas son las siguientes: “Accidente sin heridos ni víctimas fatales con bloqueo menor de pistas” (I_1p), “Accidente sin heridos ni víctimas fatales con bloqueo mayor de pistas” (I_2p), “Accidente con al menos un herido o muerto y bloqueo menor de pistas” (HM_1p) y “Accidente con al menos un herido o muerto y bloqueo mayor de pistas” (HM_2p). Las funciones de propensión estimadas para este modelo se presentan en la Tabla 4-4.

En algunas funciones presentadas en las Tablas 4-1 a 4-4 hay variables que corresponden a interacciones entre las variables originales descritas en el capítulo 3. Por ejemplo, la variable “LLUVIAxCURVA” corresponde a la interacción entre las variables LLUVIA y CURVA.

Tabla 4-4: Funciones de Propensión – Modelo 5 Alternativas.

Alt.	Razón	Variables en Función de Propensión
I_1p	1:1 – 1:5	DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA PENDIENTE EN BAJADA x ESTADO CONGESTIÓN PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA CONSTANTE I_1p
I_2p	1:1 – 1:5	DISTANCIA ENTRE ESTACIONES 1 Y 3 PENDIENTE EN SUBIDA DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA LLUVIA x CURVA PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA CONSTANTE I_2p
HM_1p	1:1 – 1:5	DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CONGESTIÓN ESTADO CUELLO DE BOTELLA DURANTE EL DÍA EN DÍA DE SEMANA PENDIENTE EN BAJADA x ESTADO FINAL DE COLA CONSTANTE HM_1p
HM_2p	1:1 – 1:5	SEMANA ENTRADA DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS ESTADO FINAL DE COLA ESTADO CUELLO DE BOTELLA CONSTANTE HM_2p

4.1.2 Signos Esperados

En esta sección se exponen los signos esperados de los parámetros asociados a cada tipo de variable. Cuando un parámetro posee un signo positivo, se interpreta como un aumento en la propensión de ocurrencia de un accidente cuando la variable asociada a ese parámetro se activa (cuando es *dummy*) o aumenta (cuando es continua). Por ende, significa un aumento en la probabilidad de ocurrencia, asumiendo que el resto de las variables permanecen fijas.

De acuerdo a las funciones de propensión presentadas en la sección anterior, nos encontramos con las siguientes variables presentes en los cuatro tipos de modelos: los Estados de Tráfico, las desviaciones estándar de la velocidad (aguas arriba y aguas abajo del lugar del accidente), las variables de geometría y las variables de condiciones climáticas.

Los Estados de Tráfico y desviaciones estándar de la velocidad se consideran como variables de política, por lo tanto, los parámetros asociados a estos factores siempre deben ser incluidos en la formulación del modelo, y por lo mismo se puede asumir *a priori* un signo esperado para los parámetros asociados a éstas. En cambio, las condiciones climáticas y geometría de la vía no son consideradas como variables de política, pero esto no significa que podrían aportar información a alguno de los modelos cuando sus parámetros son estadísticamente significativos.

Para el modelo binario, se esperan signos positivos de los parámetros asociados a las desviaciones estándar y Estados de Tráfico correspondientes a la alternativa que representa la ocurrencia de un accidente cualquiera. Por otro lado, el signo de los parámetros estadísticamente significativos asociados a las variables de geometría y condiciones climáticas, se consideran como indefinidos, ya que no necesariamente todas las variables de estos tipos se consideran como riesgosas.

Para el modelo de 3 alternativas solo según DP, se espera que ocurra lo mismo que en el modelo anterior. Es decir, cualquiera fuese el tipo de accidente, se esperan signos positivos para los parámetros asociados a las variables de tráfico, y signos indefinidos para las variables de geometría y clima.

Para los modelos de 3 alternativas según DT y 5 alternativas DP+DT, se espera que el signo de los parámetros asociados al Estado 3 sea indefinido para las alternativas “A_{2p}” e “I_{2p}”, respectivamente. Esto se debe a que un estado de tráfico congestionado no necesariamente implica riesgo de ocurrencia de un accidente con bloqueo mayor en la vía. Con respecto al resto de los parámetros de las distintas alternativas, se espera que ocurra lo mismo que en los modelos anteriores.

4.1.3 Resultados Generales de Estimación

Este apartado muestra, en primer lugar, resultados generales de estimación que indican el desempeño de todos los tipos de modelos, utilizando el 85% de cada una de las muestras definidas por las distintas razones Caso-Control.

Luego, en las secciones siguientes (4.1.4 a 4.1.7), se muestran los resultados con respecto a los parámetros estimados, junto a un análisis e interpretación de los mismos. En cada modelo este análisis se estructura, en primer lugar, con una verificación de la significancia de los parámetros, signos y magnitudes, y luego se analizan e interpretan los resultados por tipo de variable. Para cada tipo de modelo, se presentan los parámetros obtenidos a partir de la estimación realizada con la muestra más grande, es decir, con la base de datos que considera la razón Caso-Control de 1:5, ya que en el área de accidentes viales lo recomendado y usualmente usado es tener 4 o 5 Controles por Caso. El resto de los parámetros obtenidos con las muestras desde 1:1 hasta 1:4 se muestran en el Anexo E: Resultados de MNL.

Un análisis general con respecto a la estimación, se realiza utilizando los indicadores FPR, ER (valor esperado de FPR) y E(CR). En este contexto, FPR no corresponde precisamente al porcentaje correctamente predicho, sino que al porcentaje correctamente estimado. Además, se sabe que el valor esperado de CR es igual a CR, donde el modelo binario posee un 50%, ambos modelos de 3 alternativas (DP o DT) un 33%, y el modelo más complejo un 20%.

El análisis consiste en verificar que ER y FPR son similares y mayores que E(CR), puesto que al cumplirse esta condición, se puede decir que los modelos son razonables e informativos en cuanto a la estimación. Para cada modelo se muestran los indicadores calculados de ER y FPR en las Tabla 4-5 y Tabla 4-6, respectivamente.

Como se puede observar, los indicadores ER y FPR de cada modelo son muy similares y en algunos casos iguales entre sí. Esto indica que los modelos son todos razonables e informativos, ya que cada uno de los valores de cada modelo son mayores que el E(CR) correspondiente.

Tabla 4-5: Resultados Generales de Estimación – ER (%).

Bases de Datos	Binario	3 DP	3 DT	DP + DT
1:1	72,7	62,3	66,8	59,9
1:2	79,4	72,3	74,9	70,9
1:3	82,9	78,0	79,5	76,9
1:4	84,9	81,6	82,4	80,9
1:5	86,8	84,2	84,7	83,8

Tabla 4-6: Resultados Generales de Estimación – FPR (%).

Bases de Datos	Binario	3 DP	3 DT	DP + DT
1:1	72,5	62,3	67,1	59,9
1:2	77,9	72,0	74,7	71,0
1:3	81,5	77,6	79,1	77,1
1:4	83,9	81,2	82,2	80,9
1:5	86,0	83,9	84,7	83,8

También, se analiza el nivel de información que estos modelos entregan. Si miramos los resultados de FPR, ellos nos indican que el mejor modelo es el binario con la muestra de 1:5, ya que posee un 86% de éxito de estimación. Sin embargo, un modelo que considera dos respuestas entrega menos información en comparación con el resto de los modelos. Es decir, mientras más compleja es la especificación de los modelos, más información es aportada. Por su parte, los resultados muestran que cuando la complejidad de especificación aumenta se experimenta una baja en el desempeño de los modelos, pero esta disminución se contrasta con un modelo más complejo que entrega información más valiosa y desagregada.

En resumen, todos los indicadores muestran una mejora cuando la muestra crece, pero se ven comprometidos cuando la especificación se vuelve más compleja. Es decir, cuando la especificación es más desagregada, el desempeño de los modelos empeora.

A pesar que el modelo más complejo DP+DT no supera a ninguno de los otros en cuanto a FPR, cualquiera fuese la razón Caso-Control usada para estimar, este modelo es el que aporta información más desagregada, por lo que tener solo un 2,2 puntos

porcentuales de diferencia de FPR con respecto al modelo binario 1:5, no se considera como una diferencia importante. Por lo tanto, este *trade-off* entre nivel de desempeño y entrega de información es aceptado, ya que el modelo DP+DT entrega información desagregada que permite calcular la pérdida social por tener un accidente, considerando no solo el costo social de tener un accidente con heridos o víctimas fatales, sino que también el costo social de la congestión que se puede generar al causar una disrupción del tráfico a causa del bloqueo de pistas.

Antes de presentar los parámetros estimados de los cuatro tipos de modelos, usando la muestra de 1:5, es necesario mencionar otra manera que respalda el uso de una razón Caso-Control de 1:5, la cual demuestra que a medida que la muestra crece los errores estándar de los parámetros estimados disminuyen, es decir, los intervalos de confianza se van estrechando. Esta evidencia es algo mencionado en estudios anteriores y se puede encontrar en el Anexo F: Errores Estándar Estimados.

4.1.4 Parámetros Estimados del Modelo Binario

El modelo binario estimado con una muestra de 1:5 posee un indicador “FPR” de éxito de estimación igual a 86% y muestra sus resultados en la Tabla 4-7.

Tabla 4-7: Resultados de Estimación con 85% de muestra – Modelo Binario 1:5.

Parámetro	Valor	Valor-p	Riesgo Relativo
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA	0,548	0,02	1,572
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA	0,716	0,01	1,794
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	0,057	0,00	1,049
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	0,027	0,00	1,023
ESTADO FINAL DE COLA	2,080	0,00	4,699
ESTADO CONGESTIÓN	0,866	0,00	2,022
ESTADO CUELLO DE BOTELLA	1,840	0,00	4,029
DISTANCIA ENTRE ESTACIONES 1 Y 3	0,0005	0,02	1,0005
SEMANA	-0,222	0,06	0,831
CONSTANTE ACCIDENTE	-3,380	0,00	--
CONSTANTE NO-ACCIDENTE	0	--	--
Log-verosimilitud	-1681,872		

i. *Significancia, Signo y Magnitud de los Parámetros*

Los resultados de este modelo muestran que todos los parámetros son significativos sobre un 95% de confianza, pero, si se relaja esta restricción, el parámetro asociado a la variable SEMANA fue significativo al 94%.

Como se esperaba, todos los parámetros muestran magnitudes y signos esperados. Las magnitudes que se esperan están dadas por el valor de la constante y los valores que toman las variables explicativas en la muestra. Por lo tanto, se espera que cualquier parámetro estimado no supere, en valor absoluto, el valor de la constante.

ii. *Análisis de Parámetros asociados a interacciones*

Los parámetros asociados a las variables “PENDIENTE EN BAJADAxESTADO FINAL DE COLA” y “PENDIENTE EN BAJADAxESTADO CUELLO DE BOTELLA”, indican un aumento en la probabilidad de ocurrencia de un accidente en un 57,2% $((1,572-1)*100)$ y 79,4%, respectivamente. La interpretación de éstos corresponde a que en el caso de estar presente en una pendiente de bajada y las velocidades tanto aguas arriba como aguas abajo, cambien de altas a bajas o viceversa, aumenta el riesgo de ocurrencia de un accidente.

En el primer caso (“PENDIENTE EN BAJADAxESTADO FINAL DE COLA”), la probabilidad aumenta puesto que una pendiente en bajada podría volver más vulnerables a los conductores en su capacidad de frenado para responder frente a cambios de velocidades de altas a bajas, o bien, una pendiente de bajada en un cierto tramo, podría confundir la presencia de congestión en la estación de aguas abajo.

En cambio, si el cambio de velocidades fuese al revés (“PENDIENTE EN BAJADAxESTADO CUELLO DE BOTELLA”), la probabilidad aumentaría si la aceleración de los conductores podría generar un mayor descontrol del vehículo en una pendiente en bajada o entrando a ella.

iii. *Análisis de Parámetros asociados a Desviaciones Estándar de Velocidad*

Los parámetros asociados a las variables continuas de la desviación estándar de la velocidad para ambas estaciones (1 y 3) en el intervalo de tiempo más cercano a la hora

del incidente (5 minutos previos al accidente), indican que un aumento en las desviaciones estándar aumenta la probabilidad de accidente, ya que este aumento significa que los conductores deben ajustar su velocidad más frecuentemente y son más propensos a cometer errores en mantener la distancia de seguridad (Lee et al., 2002). Estos resultados se relacionan con lo expuesto por Abdel-Aty et al. (2004), mencionando que altas tasas de ocupación en la estación 1 junto con altas variaciones en la velocidad en la estación 3, aumentan la probabilidad de ocurrencia de un choque entre ellas. Sin embargo, ambos cambios dentro de una ventana desde 5 hasta 10 minutos previos al accidente.

En términos específicos, cuando las desviaciones estándar aguas arriba y aguas abajo aumentan en una unidad, manteniendo fijas las demás variables, la probabilidad de ocurrencia de un accidente aumenta en un 4,9% y 2,3% respectivamente. En Zheng et al. (2010) se habla de un aumento del 8% en la probabilidad de ocurrencia cuando la desviación estándar de la velocidad aguas arriba aumenta, lo cual **no** difiere mucho del 4,9% obtenido en este modelo. Por lo tanto, una reducción en este indicador de variación de velocidad es ventajoso, porque aumenta la seguridad reduciendo la probabilidad de accidente (Oh et al., 2005).

iv. *Análisis de Parámetros asociados a Estados de Tráfico*

En cuanto a los parámetros asociados a los Estados de Tráfico, se considera el Estado de Flujo Libre como base, asumiéndolo como el menos riesgoso. Es interesante ver las magnitudes de estos parámetros, ya que todos en conjunto indican que la presencia de ellos hace más probable un accidente en comparación al Estado base. Zheng (2012) se refiere la probabilidad de ocurrencia de accidente como *Crash Occurrence Likelihood* (COL), y en sus resultados resalta que la COL de un estado congestionado es 6 veces la COL de un Estado de Flujo Libre, y la COL en un Estado de Transición (Estados “Final de Cola” o “Cuello de Botella” en esta investigación) es 1,6 veces la COL en Flujo Libre. Sin embargo, el Estado “Final de Cola” muestra tener una probabilidad de ocurrencia de un accidente 4,7 (4,699) veces mayor que el Estado de Flujo Libre. El Estado “Congestión” es aproximadamente 2 veces más riesgoso que el

Estado “Flujo Libre”, y el Estado “Cuello de Botella” posee una COL 4 veces la COL del Estado base. Por otro lado, Yeo et al. (2013) define Estados de Tráfico, similares a los de esta investigación, muestra que la tasa de ocurrencia de accidentes para los Estados *Bottleneck* (Estado “Cuello de Botella”), *Back Of Queue* (Estado “Final de Cola”) y *Congestion* (Estado “Congestión”), son aproximadamente 5 veces mayor que para el Estado de Flujo Libre. En este modelo binario se muestra que ese valor es, en promedio, aproximadamente 4 veces mayor que para estados de Flujo Libre.

Por lo tanto, de estos resultados, se desprende que el Estado “Final de Cola” (que considera velocidades altas aguas arriba y bajas aguas abajo) es el más riesgoso de todos en comparación al Estado de Flujo Libre. Esta conclusión es igual a la reportada en Xu et al. (2012). Además, también asumen que el Estado que posee condiciones de Flujo Libre, tanto aguas arriba como aguas abajo, es el menos riesgoso.

v. *Análisis del resto de los Parámetros*

Los parámetros de las variables “DISTANCIA ENTRE ESTACIONES 1 Y 3” y SEMANA poseen signos razonables, ya que mientras más grande sea el tramo (sección de vía definida en el sección 3.2), mayor es la probabilidad de que ocurra un accidente, y, además, es menos probable tener un accidente en día de semana que en un día sábado o domingo. La variable “DISTANCIA ENTRE ESTACIONES 1 Y 3” aumenta la probabilidad de accidente en un 0,05%, cuando el resto de las variables se mantiene constante y el largo del tramo crece en una unidad. En cambio, si la variable SEMANA toma el valor 1, y asumimos *ceteris paribus* para el resto, la probabilidad disminuye en un 16,9%.

En cuanto a los parámetros asociados a las condiciones climáticas y de geometría, éstos no parecen ser significativos en este modelo, lo cual se puede respaldar por el hecho de que la vía en estudio es una autopista que carece de características desafiantes para los conductores. Estas características desafiantes se definen como pendientes muy empinadas y curvas cerradas (Ahmed et al., 2012a, 2012b).

4.1.5 Parámetros Estimados del Modelo de 3 Alternativas DP

El modelo que considera la severidad solo como DP posee un indicador “FPR” de 83,9% y sus resultados, en cuanto a la estimación, se aprecian en la Tabla 4-8.

Tabla 4-8: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:5.

Parámetro	Valor	Valor-p	Riesgo Relativo
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos)	0,716	0,01	1,928
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,000	0,00	2,376
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,051	0,00	1,042
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos)	0,068	0,00	1,060
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,030	0,00	1,026
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos)	0,024	0,01	1,019
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos)	-0,245	0,06	0,796
ENTRADA (Con Heridos)	-0,243	0,08	0,797
ESTADO FINAL DE COLA (Sin Heridos)	2,420	0,00	5,939
ESTADO FINAL DE COLA (Con Heridos)	1,910	0,00	3,566
ESTADO CONGESTIÓN (Sin Heridos)	1,040	0,00	2,390
ESTADO CONGESTIÓN (Con Heridos)	0,503	0,08	1,397
ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,780	0,00	3,742
ESTADO CUELLO DE BOTELLA (Con Heridos)	1,830	0,00	3,934
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos)	0,001	0,04	1,0005
MERGE (Con Heridos)	0,680	0,08	1,861
CONSTATE (Con Heridos)	-3,910	0,00	--
CONSTATE (Sin Heridos)	-4,080	0,00	--
CONSTATE (No Accidente)	0	--	--
Log-verosimilitud	-2193,835		

i. Significancia, Signo y Magnitud de los Parámetros

Los resultados de este modelo muestran que todos los parámetros son significativos sobre un 95% de confianza, exceptuando los parámetros de la alternativa con heridos asociados a las variables: “DURANTE EL DÍA EN DÍA DE SEMANA”, ENTRADA, “ESTADO CONGESTIÓN” y MERGE; que fueron significativos al menos al 92%. Al igual que el modelo binario, todos los parámetros muestran magnitudes y signos esperados.

ii. *Análisis de Parámetros asociados a interacciones*

Interpretando los parámetros, se observa que cuando se tiene una pendiente de bajada y un Estado de Tráfico con velocidades altas aguas arriba y bajas aguas abajo (“Final de Cola”), aumenta aproximadamente 2 veces la probabilidad de tener un accidente con al menos un herido o víctima fatal, si el resto de las variables permanece constante.

Además, si tenemos nuevamente una pendiente de bajada y un Estado del tipo cuello de botella, aumenta 2,4 veces la probabilidad de tener un accidente menos severo (sin heridos ni muertos). Bham et al. (2012) infieren que la presencia de gradiente aumenta 2,25 veces la severidad de un accidente comparado con secciones planas.

iii. *Análisis de Parámetros asociados a Desviaciones Estándar de Velocidad*

Por otro lado, para ambas alternativas de accidente aumenta la probabilidad de ocurrencia cuando las desviaciones estándar aumentan tanto en la estación de aguas arriba como en la de aguas abajo.

Según los resultados, se infiere que al aumentar en una unidad la desviación estándar de las velocidades de aguas arriba, manteniendo los demás parámetros fijos, aumenta en un 4,2% la probabilidad de ocurrencia de un accidente que resulta solo en daños materiales. Por su parte, en el mismo escenario, aumenta en un 6% la probabilidad de ocurrencia de un accidente con al menos un herido o muerto.

En cambio, cuando aumenta en una unidad la desviación estándar en la estación de aguas abajo, las probabilidades de ocurrencia de un accidente leve y uno severo aumentan en 3% y 2,4%, respectivamente.

iv. *Análisis de Parámetros asociados a Estados de Tráfico*

En cuanto a los parámetros estimados asociados con los Estados de Tráfico, también se puede concluir que los parámetros de los Estados considerados *a priori* más riesgosos, tienen valores mayores que los demás, como los Estados de Tráfico 1 y 4, al igual que en el modelo anterior.

Para un accidente leve, en presencia del Estado “Final de Cola”, la probabilidad de ocurrencia es aproximadamente 6 veces mayor con respecto al Estado “Flujo Libre”, manteniendo el resto de las variables constante. En cambio, si el Estado “Congestión” estuviese activo, la probabilidad es 2,3 veces mayor, y 3,7 veces mayor en presencia del Estado “Cuello de Botella”.

Para accidentes severos, la presencia de los Estados “Final de Cola”, “Congestión” y “Cuello de Botella” poseen probabilidades de ocurrencia 3,6, 1,4 y 3,9 veces mayor con respecto al Estado base, respectivamente.

Por lo tanto, se infiere que la presencia de los Estados “Final de Cola” y “Congestión” con respecto al Estado “Flujo Libre”, aumentan en mayor magnitud la probabilidad de ocurrencia de un accidente leve en comparación con un accidente severo. Es decir, los riesgos relativos ya sea del Estado “Final de Cola” o “Congestión” son mayores para un accidente leve que para uno severo. Por su parte, la presencia del Estado “Cuello de Botella” genera un aumento similar en la ocurrencia de cualquiera de los dos tipos de accidentes.

Además, se concluye que los estados de transición (“Final de Cola” y “Cuello de Botella”), en general, generan más riesgos de ocurrencia de un accidente cualquiera (leve o severo).

Estudios que consideran factores agregados de condiciones de tráfico y no Estados de Tráfico propiamente tal, no encontraron evidencia suficiente que explique la ocurrencia de accidentes a causa del tráfico. En Quddus et al. (2010) se quiso investigar el efecto de la congestión en la severidad de accidentes usando la demora total como *proxi*, y Yang et al. (2011) usaron el tráfico promedio diario tanto en rampas de entrada y salida como en la vía principal.

Por otro lado, Haleem & Abdel-Aty (2010) mencionan que la probabilidad de mayor severidad según lesión (y fatalidad) está siempre asociada con una reducción en el AADT (*Annual Average Daily Traffic*), así como también un aumento en los límites de velocidad.

Es interesante mencionar lo que se ha descubierto con respecto a la velocidad y los límites impuestos en las vías estudiadas en la literatura revisada, puesto que la velocidad

es la variable utilizada en esta investigación para definir los Estados de Tráfico. Dentro de los hallazgos más relevantes, Abdel-Aty (2003) menciona que un aumento en la razón entre la velocidad estimada a la hora del accidente y el límite impuesto, aumenta la probabilidad de lesiones severas. Eluru & Bhat (2007) dicen que la lesión del conductor tiende a ser más severa cuando existen límites de velocidad de categoría media (42 y 103 kph, respectivamente), y que la lesión tiende a ser más grave cuando se superan los límites altos, con respecto a la severidad observada de la lesión cuando existen límites menores que la categoría media. Por su parte, Yang et al. (2011) encontraron que un aumento en la diferencia entre el límite de velocidad impuesto en la vía principal y límite de una rampa aumenta la severidad de un choque.

v. *Análisis de Parámetros asociados a Variables Geométricas*

El valor estimado del parámetro asociado a la variable ENTRADA disminuye la propensión de un accidente severo. Su riesgo relativo promedio indica que cuando esta variable se activa (toma el valor 1) la probabilidad de ocurrencia de accidentes severos disminuye en un 20,3%, si el resto de las variables permanece constante. Es decir, cuando un tramo posee una pista de entrada a la autopista los conductores tienden a conducir de manera más precavida al percatarse de la presencia de ésta.

Por otro lado, el parámetro estimado asociado a la variable “DISTANCIA ENTRE ESTACIONES 1 Y 3” aumenta la propensión de accidentes leves. Esto se debe a que, al tener tramos más largos, es más probable que ocurra un accidente leve. Dicha probabilidad aumenta en un 0,05% cuando el largo del tramo crece un metro y el resto de las variables permanece constante. Esto se explica con la siguiente analogía: siuviésemos que lanzar una pelota hacia el agujero de una urna que se encuentra a una distancia muy lejana, la probabilidad de que la pelota caiga dentro de ella es muy baja, pero ésta tiende a aumentar cuando el agujero aumenta en tamaño y la urna permanece en el lugar original.

En cuanto al parámetro asociado a la variable MERGE, se puede decir que cuando se está presente en un tramo en donde confluyen los dos ejes importantes de la autopista (General Velásquez y Ruta 5), la probabilidad de ocurrencia de un accidente severo

aumenta en un 81,6%, considerando *ceteris paribus* para el resto de las variables con parámetros significativos. La confluencia de ejes genera un aumento en la cantidad de pistas del tramo en análisis, y según Yang et al. (2011) este aumento genera una mayor probabilidad de accidentes severos.

vi. *Análisis del resto de los Parámetros*

El parámetro asociado a la variable “DURANTE EL DÍA EN DÍA DE SEMANA” disminuye la propensión de un accidente severo. La activación de esta variable (es decir, cuando es igual a 1) indica una disminución del 20,4% en la probabilidad de ocurrencia de accidentes severos, si el resto de las variables permanecen fijas. En términos cualitativos, es menos probable observar un accidente severo durante un día de semana en horario diurno, puesto que al tener luz de día y estar en un día de semana las condiciones de conducción pueden ser mejores, ya sea por iluminación favorable, o porque en la noche y fin de semana el flujo es menor, lo que se traduce en que los conductores manejen a velocidades mayores, lo cual es directamente proporcional con el grado de severidad.

Sin embargo, esto no se puede considerar como algo concluyente, puesto que existe muchas otras combinaciones para el resto de las variables. Eluru & Bhat (2007) dicen que los factores ambientales indican que los accidentes que ocurren durante el día (6am – 7pm) tienden a ser menos severos que aquellos ocurridos en otras horas, lo que se puede deber a la existencia de menos flujo fuera de ese horario.

4.1.6 Parámetros Estimados del Modelo de 3 Alternativas DT

El modelo que considera como severidad solo la disrupción en el tráfico, en cuanto al porcentaje correctamente estimado, muestra un desempeño del 84,7% y los resultados de los parámetros estimados se aprecian en la Tabla 4-9.

Tabla 4-9: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:5.

Parámetro	Valor	Valor-p	Riesgo Relativo
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Bloqueo Menor)	-0,807	0,01	0,487
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	0,850	0,00	2,039
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,048	0,00	1,038
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,101	0,00	1,095
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,023	0,00	1,018
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,059	0,00	1,055
ESTADO FINAL DE COLA (Bloqueo Menor)	2,520	0,00	6,142
ESTADO FINAL DE COLA (Bloqueo Mayor)	1,680	0,00	2,652
ESTADO CONGESTIÓN (Bloqueo Menor)	1,600	0,00	3,687
ESTADO CONGESTIÓN (Bloqueo Mayor)	-1,130	0,02	0,240
ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	1,780	0,00	3,806
ESTADO CUELLO DE BOTELLA (Bloqueo Mayor)	1,680	0,00	3,444
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Bloqueo Mayor)	0,001	0,01	1,001
LLUVIA x CURVA (Bloqueo Mayor)	0,788	0,12	2,112
PENDIENTE EN SUBIDA (Bloqueo Mayor)	-0,305	0,07	0,746
SEMANA (Bloqueo Mayor)	-0,340	0,08	0,722
CONSTANTE (Bloqueo Menor)	-3,320	0,00	--
CONSTANTE (Bloqueo Mayor)	-5,720	0,00	--
CONSTANTE (No Accidente)	0	--	--
Log-verosimilitud	-2068,407		

i. *Significancia, Signo y Magnitud de los Parámetros*

En general, los resultados de este modelo muestran que los parámetros son significativos por sobre un nivel de confianza del 95%, exceptuando los parámetros de la alternativa “Bloqueo Mayor” asociados a las variables: “LLUVIAxCURVA”, “PENDIENTE EN SUBIDA” y SEMANA; que fueron significativos al menos al 88% de confianza. Al igual que los modelos presentados, se obtienen parámetros con signos y magnitudes esperados.

ii. *Análisis de Parámetros asociados a interacciones*

Analizando los parámetros asociados a interacciones entre variables, se observa que cuando se tiene una pendiente en bajada y el Estado “Congestión”, disminuye en un 51,3% la probabilidad de ocurrencia de un accidente con poca disrupción del tráfico

(A_1p). En cambio, si fuese el Estado “Cuello de Botella” la probabilidad de la misma alternativa sería aproximadamente 2 veces mayor que si no existiese pendiente de bajada y/o el Estado “Cuello de Botella”.

Por otro lado, el parámetro asociado a la variable “LLUVIAxCURVA” indica que cuando el tramo experimenta lluvia y además presenta curvas, la probabilidad de sufrir un accidente con “Bloqueo Mayor” es 2,1 veces mayor que si no estuviese lloviendo y/o el tramo no presentase curvas. Esto se traduce en que cuando el tramo presenta condiciones adversas de clima y geometría para los conductores, que podrían eventualmente sufrir un accidente de este tipo.

iii. Análisis de Parámetros asociados a Desviaciones Estándar de Velocidad

De acuerdo a estos parámetros, se observa que los parámetros de las desviaciones estándar de las velocidades tanto aguas arriba como aguas abajo, aumentan la propensión de un accidente, cualquiera fuese el tipo.

A partir de los resultados se desprende que (al mantener el resto de los parámetros fijos), un aumento de una unidad en la desviación estándar de las velocidades de aguas arriba incrementa en un 3,8%, la probabilidad de ocurrencia de un accidente con “Bloqueo Menor”. Por su parte, en el mismo escenario, el aumento de la probabilidad de ocurrencia de un accidente con “Bloqueo Mayor” es 9,5%.

En cambio, cuando aumenta en una unidad la desviación estándar en la estación de aguas abajo, las probabilidades de ocurrencia de un accidente con “Bloqueo Menor” y “Bloqueo Mayor” aumentan en un 1,8% y 5,5%, respectivamente.

iv. Análisis de Parámetros asociados a Estados de Tráfico

Al igual que en los modelos anteriores, se vuelve a confirmar que la presencia de Estados de Tráfico de transición (“Final de Cola” y “Cuello de Botella”) son los que presentan mayor riesgo de accidente, con respecto a un Estado de Flujo Libre. Para un accidente con “Bloqueo Menor” en presencia de los Estados “Final de Cola”, “Congestión” y “Cuello de Botella”, la probabilidad de ocurrencia es 6,1, 3,7 y 3,8 veces mayor con respecto al Estado “Flujo Libre”, respectivamente.

Para un accidente con bloqueo mayor, la probabilidad de ocurrencia es 2,6 y 4,4 veces mayor con respecto al Estado “Flujo Libre”, en presencia de los Estados “Final de Cola” y “Cuello de Botella”, respectivamente. En cambio, la probabilidad disminuye en un 76% con respecto al Estado “Flujo Libre” cuando se está en presencia del Estado “Congestión”. Esto indica que en un estado congestionado es menos probable tener un accidente que provoque bloqueo de dos o más pistas, debido a las velocidades bajas, donde es más probable la ocurrencia de accidente con bloqueo menor.

v. *Análisis de Parámetros asociados a Variables Geométricas*

El parámetro asociado a la variable “DISTANCIA ENTRE ESTACIONES 1 Y 3” aumenta el riesgo de ocurrencia de un accidente con bloqueo mayor de la vía. Esto se debe a que, a tener tramos cada vez más largos, es más probable que ocurra un accidente con alto bloqueo de pistas. Los datos indican que dicha probabilidad aumenta en un 0,1% cuando el largo del tramo se incrementa en un metro y el resto de las variables permanece fijo, pero esto no se considera como concluyente debido a que su interpretación no es clara.

En cuanto a la variable “PENDIENTE EN SUBIDA”, su parámetro indica una disminución en la propensión de un accidente tipo “Bloqueo Mayor”. El riesgo relativo de este parámetro indica que cuando la variable se encuentra activa esta probabilidad disminuye en un 25,4%.

vi. *Análisis del resto de los Parámetros*

El parámetro de la variable SEMANA indica una disminución en el riesgo de ocurrencia de un accidente con bloqueo mayor. La interpretación de este parámetro dice que es menos probable observar un accidente de este tipo durante un día laboral. En detalle, la probabilidad disminuye en un 27,8% cuando es un día de semana.

4.1.7 Parámetros Estimados del Modelo de 5 Alternativas

Este modelo combina los efectos de los dos modelos anteriores, es decir, considera la severidad de accidentes según DP y DT.

El porcentaje correctamente estimado, es decir, el nivel de desempeño de este modelo es de un 83,8%; y sus resultados se muestran en la Tabla 4-10.

i. *Significancia, Signo y Magnitud de los Parámetros*

En términos generales, los resultados muestran que los parámetros, al menos la mayoría de ellos, son significativos por sobre un nivel de confianza del 95%. Aquellos que no están sobre el 95% son al menos significativos al 89% de confianza, de los cuales se tienen los siguientes parámetros:

- El parámetro de la alternativa “Sin Heridos y Bloqueo Mayor” asociado a la variable “PENDIETE EN BAJADAxESTADO CUELLO DE BOTELLA”.
- El parámetro de la alternativa “Con Heridos y Bloqueo Menor” asociado a la variable “DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS”.
- Y el parámetro del accidente “Con Heridos y Bloqueo Mayor” asociado a la variable SEMANA.

Todos los parámetros poseen magnitudes y signos esperados.

Cabe destacar que el parámetro asociado al Estado “Congestión” de la alternativa “Con Heridos y Bloqueo Mayor” no aparece en el modelo, por presentar un valor-p mayor 0,2, lo cual indica un nivel de confianza menor al 80% aceptado. Además, el comportamiento del Estado “Congestión” cuando se presenta una alternativa con bloqueo mayor, no se considera como una variable de política, ya que, *a priori* no es claro el efecto que ésta pueda producir en este tipo de accidentes.

Tabla 4-10: Resultados de Estimación con 85% de muestra – Modelo DP + DT 1:5.

Parámetro	Valor	Valor-p	Riesgo Relativo
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	0,714	0,03	1,961
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,770	0,02	0,488
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,080	0,00	2,537
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,090	0,07	2,562
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,047	0,00	1,038
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,086	0,00	1,079
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,047	0,00	1,038
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,104	0,00	1,099
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,024	0,00	1,019
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,095	0,00	1,094
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,018	0,08	1,013
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,029	0,03	1,025
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,363	0,02	0,708
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,470	0,04	0,631
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,570	0,00	6,759
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	1,060	0,01	1,493
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	2,030	0,00	3,939
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	2,110	0,00	4,267
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,780	0,00	4,495
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-2,790	0,02	0,047
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,821	0,01	1,723
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,730	0,00	3,594
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,450	0,01	2,716
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	2,060	0,00	4,999
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,410	0,00	2,609
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,002	0,00	1,002
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	1,450	0,02	4,080
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,825	0,00	0,444
SEMANA (Con Heridos y Bloqueo Mayor)	-0,390	0,11	0,684
CONSTANTE (Con Heridos y Bloqueo Menor)	-4,020	0,00	--
CONSTANTE (Con Heridos y Bloqueo Mayor)	-5,270	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Menor)	-3,800	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-7,550	0,00	--
CONSTANTE (No Accidente)	0	--	--
Log-verosimilitud	-2552,638		

ii. *Análisis de Parámetros asociados a interacciones*

En específico, los resultados del modelo de especificación más desagregada muestran, en primera instancia, que cuando se tiene pendiente en bajada y el Estado “Final de Cola”, la probabilidad de ocurrencia un accidente “Con Heridos y Bloqueo Menor” es 2 veces mayor que si no existiese pendiente de bajada y/o el Estado “Final de Cola”. Esto, de alguna manera, se respalda en Bham et al. (2012), mencionando que la presencia de gradiente aumenta la severidad de un choque 2,25 veces en comparación con secciones niveladas.

Manteniendo la pendiente de bajada, pero ahora con un Estado “Congestión”, la probabilidad de ocurrencia de un accidente “Sin Heridos y Bloqueo Menor” disminuye a 0,49 veces la probabilidad cuando no existe pendiente de bajada y/o el Estado “Congestión”. En cambio, si el estado fuese el de “Cuello de Botella”, manteniendo el mismo escenario de pendiente de bajada activa, tanto la probabilidad de ocurrencia del mismo tipo de accidente como del “Sin Heridos y Bloqueo Mayor”, ambas se vuelven 2,5 veces mayor que si no existiese pendiente de bajada y/o el Estado “Cuello de Botella”.

El valor estimado para el parámetro asociado a la variable “LLUVIAxCURVA” de la alternativa “Sin Heridos y Bloqueo Mayor” indica un aumento en la propensión de ocurrencia de ésta. Si la variable “LLUVIAxCURVA” se activa (es igual a 1) y el resto de las variables permanecen fijas, la probabilidad de “Sin Heridos y Bloqueo Mayor” sería 4 veces mayor que si no estuvieses lloviendo y/o el tramo no presentase curvas.

En la literatura algunos mencionan que indicadores de estación del año son significativos, implicando que las condiciones climáticas pueden afectar la severidad de un accidente (Jung et al., 2010). Por un lado, Yang et al. (2011) dicen que una buena iluminación y condiciones favorables del clima disminuyen la probabilidad de ocurrencia de un accidente severo, mientras que Eluru & Bhat (2007) mencionan que tener condiciones pobres de iluminación y adversas del clima indican lesiones poco severas, puesto que los conductores están más alerta.

Por el lado de los factores geométricos, estudios previos, en el marco de modelos de elección discreta, mencionan que la presencia de curvas en una sección de la vía, aumenta la probabilidad de tener accidentes más perjudiciales capaces de producir lesiones (Abdel-Aty, 2003; Khattak et al., 2002). Bham et al. (2012), específicamente, menciona que la severidad aumenta 2,29 veces cuando la ocurrencia de un choque fue en presencia de una curva en comparación con una sección recta.

iii. Análisis de Parámetros asociados a Desviaciones Estándar de Velocidad

Los parámetros estimados correspondientes a las desviaciones estándar de las velocidades tanto de aguas arriba como de aguas abajo, aumentan la propensión de ocurrencia de cualquiera de los tipos de accidente. A partir de esto, se observa que un aumento de una unidad en las desviaciones de aguas arriba (manteniendo fijas el resto de las variables), la probabilidad de ocurrencia de los accidentes “Sin Heridos y Bloqueo Menor”, “Sin Heridos y Bloqueo Mayor”, “Con Heridos y Bloqueo Menor” y “Con Heridos y Bloqueo Mayor”; aumentan aproximadamente en 3,8%, 7,9%, 3,8% y 9,9%, respectivamente.

Por otra parte, si el aumento fuese el mismo, pero para las desviaciones estándar de la estación de aguas abajo, las probabilidades de ocurrencia de los accidentes: “Sin Heridos y Bloqueo Menor”, “Sin Heridos y Bloqueo Mayor”, “Con Heridos y Bloqueo Menor” y “Con Heridos y Bloqueo Mayor”; aumentan en 1,9%, 9,4%, 1,3% y 2,5%, respectivamente.

iv. Análisis de Parámetros asociados a Estados de Tráfico

La estimación de los parámetros de los Estados de Tráfico, vuelven a indicar, en términos generales, que los Estados “Final de Cola” y “Cuello de Botella” son los más riesgosos con respecto al estado base.

Específicamente, los Estados de Tráfico afectan las probabilidades de la siguiente manera:

- Para un accidente “Sin Heridos y Bloqueo Menor” con los Estados “Final de Cola”, “Congestión” o “Cuello de Botella” presentes, las probabilidades son

6,8, 4,5 y 3,6 veces mayor con respecto al Estado “Flujo Libre” (base), respectivamente.

- Para un accidente “Sin Heridos y Bloqueo Mayor” con los Estados “Final de Cola”, “Congestión” o “Cuello de Botella” iguales a 1, las probabilidades son 1,5, 0,05 y 2,7 veces mayor que la probabilidad de si el Estado base estuviese activo.
- Para un accidente “Con Heridos y Bloqueo Menor” con los Estados “Final de Cola”, “Congestión” o “Cuello de Botella” presentes, las probabilidades son 3,9, 1,7 y 4,99 veces mayor con respecto al Estado “Flujo Libre”, respectivamente.
- Para un accidente “Con Heridos y Bloqueo Mayor” con los Estados “Final de Cola” o “Cuello de Botella” activados (iguales a 1), las probabilidades son 4,3 y 2,6 veces mayores que las probabilidades cuando los Estados “Flujo Libre” o “Congestión” son iguales a 1, respectivamente. En este caso, el Estado “Congestión” no pareció comportarse significativamente diferente del Estado “Flujo Libre” base.

v. *Análisis de Parámetros asociados a Variables Geométricas*

La estimación del parámetro de la alternativa “Con Heridos y Bloqueo Mayor” asociado a la variable ENTRADA, indica una disminución en la propensión de esta alternativa. Al estar ENTRADA activa, la probabilidad de ocurrencia de un accidente “Con Heridos y Bloqueo Mayor” disminuye en un 36,9%. Esto se explica por el simple hecho de que los conductores en la vía, al percatarse de la presencia de una pista de entrada, se concentran más y manejan de manera más precavida.

Con respecto a la variable “DISTANCIA ENTRE ESTACIONES 1 Y 3”, su parámetro indica un aumento en la propensión de ocurrencia de la alternativa “Sin Heridos y Bloqueo Mayor”, lo cual es respaldado por los resultados obtenidos en los dos modelos anteriores. En detalle, cuando la variable “DISTANCIA ENTRE ESTACIONES 1 Y 3” aumenta en una unidad, la probabilidad de ocurrencia de un accidente “Sin Heridos y Bloqueo Mayor” aumenta en 0,2%.

Por otro lado, si seguimos analizando la alternativa “Sin Heridos y Bloqueo Mayor”, el parámetro asociado a la variable “PENDIENTE EN SUBIDA” dicta una disminución en el riesgo de ocurrencia de ésta. La presencia de una pendiente en subida disminuye la probabilidad de “Sin Heridos y Bloqueo Mayor” en un 55,6%, si consideramos *certeris paribus* para el resto de las variables.

vi. *Análisis del resto de los Parámetros*

El parámetro estimado, específico de “Con Heridos y Bloqueo Menor”, asociado a “DURANTE EL DÍA EN DÍA DE SEMANA”, indica que este tipo de accidentes reduce su riesgo cuando se está en un día de semana y es de día, es decir, disminuye su probabilidad de ocurrencia en un 29,2%. Lo cual es lógico, puesto que al estar desactivada esta variable, o sea, es de noche y/o es fin de semana, se tiende a circular a velocidades mayores por tener menos flujo en la vía (Eluru & Bhat, 2007), lo cual está directamente relacionado con la severidad según la dimensión de daños personales.

Por último, la estimación del parámetro de la alternativa “Con Heridos y Bloqueo Mayor” asociado a la variable SEMANA, indica una disminución en la propensión de esta alternativa. Si SEMANA fuese igual a 1, la probabilidad de ocurrencia de un accidente “Con Heridos y Bloqueo Menor” disminuye en un 31,6%. Por lo tanto, se cree que en días laborales, la velocidad es menor que en los fines de semana, generando menos posibilidad de severidad, tanto en daños personales como en disrupción del flujo vehicular.

4.1.8 Análisis Comparativo de Modelos

En esta sección se presenta una comparación entre modelos. En primera instancia se compara el modelo DP+DT con el modelo 3 DP, y luego el modelo DP+DT con el modelo 3 DT. La comparación se basa en los riesgos relativos obtenidos con los parámetros comunes entre cada modelo.

Los riesgos relativos de los parámetros comunes entre los modelos DP+DT y 3 DP se presentan en la Tabla 4-11.

Tabla 4-11: Riesgos Relativos Modelo 3 DP vs Modelo DP + DT

Parámetro	3 DP	DP + DT
DURANTE EL DÍA EN DÍA DE SEMANA	Con Heridos 0,796	Con Heridos y Bloqueo Menor 0,708
DISTANCIA ENTRE ESTACIONES 1 Y 3	Sin Heridos 1,0005	Sin Heridos y Bloqueo Mayor 1,002
ENTRADA	Con Heridos 0,797	Con Heridos y Bloqueo Mayor 0,631
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA	Con Heridos 1,928	Con Heridos y Bloqueo Menor 1,961
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA	Sin Heridos 2,376	Sin Heridos y Bloqueo Menor 2,537
		Sin Heridos y Bloqueo Mayor 2,562
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	Sin Heridos 1,042	Sin Heridos y Bloqueo Menor 1,038
		Sin Heridos y Bloqueo Mayor 1,079
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	Con Heridos 1,060	Con Heridos y Bloqueo Menor 1,038
		Con Heridos y Bloqueo Mayor 1,099
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	Sin Heridos 1,026	Sin Heridos y Bloqueo Menor 1,019
		Sin Heridos y Bloqueo Mayor 1,094
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	Con Heridos 1,019	Con Heridos y Bloqueo Menor 1,013
		Con Heridos y Bloqueo Mayor 1,025
ESTADO FINAL DE COLA	Sin Heridos 5,939	Sin Heridos y Bloqueo Menor 6,759
		Sin Heridos y Bloqueo Mayor 1,493
ESTADO FINAL DE COLA	Con Heridos 3,566	Con Heridos y Bloqueo Menor 3,939
		Con Heridos y Bloqueo Mayor 4,267
ESTADO CONGESTIÓN	Sin Heridos 2,390	Sin Heridos y Bloqueo Menor 4,495
		Sin Heridos y Bloqueo Mayor 0,047
ESTADO CONGESTIÓN	Con Heridos 1,397	Con Heridos y Bloqueo Menor 1,723
		Con Heridos y Bloqueo Mayor --
ESTADO CUELLO DE BOTELLA	Sin Heridos 3,742	Sin Heridos y Bloqueo Menor 3,594
		Sin Heridos y Bloqueo Mayor 2,716
ESTADO CUELLO DE BOTELLA	Con Heridos 3,934	Con Heridos y Bloqueo Menor 4,999
		Con Heridos y Bloqueo Mayor 2,609

Se sabe que dada la especificación de ambos modelos, los parámetros presentes en la función de propensión de una alternativa en el modelo 3 DP asociados a una variable cualquiera, deberían tener en su contraparte dos parámetros asociados a la misma variable. Por ejemplo, “DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS” en la función de la alternativa “Sin Heridos” en el modelo 3 DP,

debería ser comparado contra los parámetros asociados a la misma variable pero de las alternativas “Sin Heridos y Bloqueo Menor” e “Sin Heridos y Bloqueo Mayor” en el modelo DP+DT.

Si bien lo anterior no es algo que ocurra para todos los parámetros, no es algo que debiese ser necesariamente de esa manera. Por ejemplo, el parámetro asociado a “DURANTE EL DÍA EN DÍA DE SEMANA” está presente en la alternativa “Con Heridos” y “Con Heridos y Bloqueo Menor”, pero no en la alternativa “Con Heridos y Bloqueo Mayor”.

Lo importante a destacar en esta comparación, es que la mayoría de los riesgos relativos, entre un modelo y otro, se comportan de manera esperada. Es decir, en los casos que no ocurren lo descrito en el párrafo anterior, los riesgos relativos de un parámetro específico del modelo 3 DP se encuentra en la mitad entre los riesgos relativos de los parámetros en el modelo DP+DT. Por ejemplo, el riesgo relativo del parámetro “DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS” en la alternativa “Sin Heridos” es 1,042, y se encuentra entre 1,038 y 1,079 que son los riesgos relativos de del mismo parámetro para las alternativas “Sin Heridos y Bloqueo Menor” e “Sin Heridos y Bloqueo Mayor”, respectivamente.

Cabe destacar que cuando el riesgo relativo de un parámetro en el modelo 3 DP no es similar, o bien, no se encuentra entre los riesgos relativos de los parámetros en la contraparte, se puede deber a que el modelo DP+DT trata de explicar un fenómeno extra que el modelo 3 DP no puede, debido a la especificación.

Por otro lado, los riesgos relativos de los parámetros comunes entre los modelos DP+DT y 3 DT se presentan en la Tabla 4-12. Como era de esperar, en términos generales, en la tabla comparativa ocurre lo mismo que en la comparación anterior. Esto es lógico, ya que si se piensa que una de las alternativas en los modelos de 3 alternativas es dividida en dos alternativas en el modelo DP+DT, el efecto que en el modelo de 3 alternativas se estimó, debería dividirse en las alternativas del modelo DP+DT, y es por esto que los riesgos relativos de los modelos de 3 alternativas, en la mayoría de los casos, se encuentran entre los riesgos relativos del modelo DP+DT.

Tabla 4-12: Riesgos Relativos Modelo 3 DT vs Modelo DP + DT

Parámetro	3 DT	DP + DT
SEMANA	Bloqueo Mayor 0,722	Con Heridos y Bloqueo Mayor 0,684
DISTANCIA ENTRE ESTACIONES 1 Y 3	Bloqueo Mayor 1,001	Sin Heridos y Bloqueo Mayor 1,002
PENDIENTE EN SUBIDA	Bloqueo Mayor 0,746	Sin Heridos y Bloqueo Mayor 0,444
LLUVIA x CURVA	Bloqueo Mayor 2,112	Sin Heridos y Bloqueo Mayor 4,080
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN	Bloqueo Menor 0,487	Sin Heridos y Bloqueo Menor 0,488
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA	Bloqueo Menor 2,039	Sin Heridos y Bloqueo Menor 2,537
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	Bloqueo Menor 1,038	Sin Heridos y Bloqueo Menor 1,038
		Con Heridos y Bloqueo Menor 1,038
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	Bloqueo Mayor 1,095	Sin Heridos y Bloqueo Mayor 1,079
		Con Heridos y Bloqueo Mayor 1,099
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	Bloqueo Menor 1,018	Sin Heridos y Bloqueo Menor 1,019
		Con Heridos y Bloqueo Menor 1,013
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	Bloqueo Mayor 1,055	Sin Heridos y Bloqueo Mayor 1,094
		Con Heridos y Bloqueo Mayor 1,025
ESTADO FINAL DE COLA	Bloqueo Menor 6,142	Sin Heridos y Bloqueo Menor 6,759
		Con Heridos y Bloqueo Menor 3,939
ESTADO FINAL DE COLA	Bloqueo Mayor 2,652	Sin Heridos y Bloqueo Mayor 1,493
		Con Heridos y Bloqueo Mayor 4,267
ESTADO CONGESTIÓN	Bloqueo Menor 3,687	Sin Heridos y Bloqueo Menor 4,495
		Con Heridos y Bloqueo Menor 1,723
ESTADO CONGESTIÓN	Bloqueo Mayor 0,240	Sin Heridos y Bloqueo Mayor 0,047
		Con Heridos y Bloqueo Mayor -
ESTADO CUELLO DE BOTELLA	Bloqueo Menor 3,806	Sin Heridos y Bloqueo Menor 3,594
		Con Heridos y Bloqueo Menor 4,999
ESTADO CUELLO DE BOTELLA	Bloqueo Mayor 3,444	Sin Heridos y Bloqueo Mayor 2,716
		Con Heridos y Bloqueo Mayor 2,609

Finalmente, es necesario destacar que todas las variables que el modelo 3 DP tiene y el 3 DT no tiene (y viceversa), sí las tiene el modelo DP+DT. Ésta es otra forma de justificar el uso de una especificación más compleja, que explica los dos fenómenos DP y DT en un solo modelo.

4.1.9 Validación de los Modelos

Esta sección contiene el análisis realizado para la validación de los modelos, y un criterio que define el mejor de los modelos presentados en las secciones anteriores. El modelo escogido será comparado con los modelos presentados en la Sección 4.2.

Para realizar la validación de los modelos presentados en la subsección anterior, se utilizó el 15% restante de la muestra.

En términos generales, el análisis con respecto a la validación, se realiza utilizando los indicadores FPR, ER (valor esperado de FPR) y E(CR), al igual que en la sección anterior. En este contexto, se sabe que el valor esperado de CR es igual a CR, donde el modelo binario posee un 50%, ambos modelos de 3 alternativas (DP o DT) un 33%, y el modelo más complejo un 20%.

El análisis consiste en verificar que ER y FPR son similares y mayores que E(CR), puesto que al cumplirse esta condición, se podría decir que los modelos son razonables e informativos. Para cada modelo se muestran los indicadores calculados de FPR y ER en la Tabla 4-13.

Tabla 4-13: Resultados Validación – FPR y ER (%).

Bases de Datos	Binario		3 DP		3 DT		DP + DT	
	FPR	ER	FPR	ER	FPR	ER	FPR	ER
1:1	71,2	73,1	59,5	61,7	66,1	66,9	59,5	59,7
1:2	77,6	77,4	70,1	72,0	74,2	75,1	70,1	70,6
1:3	81,4	81,1	76,1	77,5	79,2	79,2	75,9	76,5
1:4	83,6	83,6	80,0	81,4	81,8	82,2	80,4	80,8
1:5	86,0	85,3	83,1	84,0	83,7	84,5	83,0	83,5

Como se puede observar, los indicadores ER y FPR de cada modelo son muy similares. Esto indica que los modelos, efectivamente, son todos razonables e informativos, ya que cada uno de los valores de cada modelo son mayores que el E(CR) correspondiente.

Como se mencionó en la sección anterior, un modelo con especificación más compleja entrega más información, y si miramos los resultados de FPR, ellos nos

indican que el mejor modelo, en términos de predicción, es el binario con la muestra de 1:5, ya que posee un 86% de FPR. Sin embargo, este modelo, a pesar de ser el más sencillo, entrega información de manera muy agregada.

En la Figura 4-1, se aprecia cómo los resultados cambian al considerar los dos ejes de análisis: Aumento de muestra y complejidad de especificación. Un aumento en la complejidad de especificación repercute en una disminución en la predicción de los modelos, pero esta caída, en cada modelo, se contrasta con un modelo más complejo que entrega información más valiosa y desagregada.

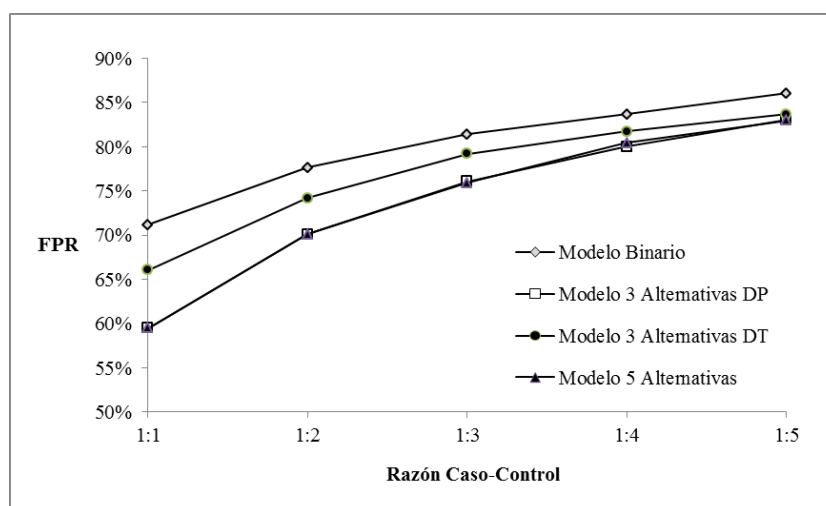


Figura 4-1: Evolución de FPR – Validación.

En resumen, si consideramos los dos ejes de análisis mencionados en el párrafo anterior para analizar la validación, se puede decir que todos los indicadores muestran una mejora cuando la muestra crece, pero se ven comprometidos cuando la especificación se vuelve más compleja, es decir, cuando la especificación es más desagregada la predicción de los modelos empeora.

En las sección de los resultados de la estimación, se cree que el mejor modelo es DP+DT con la muestra de 1:5, puesto que, a pesar de tener un poco menor desempeño que el modelo binario, entrega información desagregada más valiosa. En este caso, la validación corrobora lo esperado y la diferencia en éxito predictivo es menor (3 puntos

porcentuales), es decir, se acepta el *trade-off* entre nivel de predicción y entrega de información.

Finalmente, se concluye que la utilización de la metodología Caso-Control con MNLM tiene netamente un fin exploratorio, de forma tal de verificar que un modelo con una especificación compleja es capaz de entregar resultados lo suficientemente satisfactorios como para asegurar que la utilización de dicho tipo de especificación apunta en la dirección correcta.

4.2 Logit Mixto (MLM)

Dado los resultados y análisis de los modelos anteriores, que tienen un fin netamente exploratorio, es necesario tratar las limitaciones que posee el MNLM.

En primer lugar, se recuerda que el MNLM es un modelo que considera una escala nominal para la variable dependiente, y además, asume independencia de las alternativas, lo cual no se adapta correctamente a los datos utilizados. Si bien esto indica que una escala nominal no es lo correcto, una escala ordinal tampoco lo es, ya que los modelos ordinales imponen una restricción que no permite el crecimiento simultáneo de una variable cualquiera para distintas categorías de severidad (Malyshkina & Mannering, 2008).

En esta investigación se poseen dos dimensiones de ordenamiento de la variable dependiente, dictadas por la definición de severidad según DP y DT, donde cada una posee solo dos niveles, lo cual no justifica el uso de un modelo ordinal bivariado. Entonces, para considerar la posible correlación entre las alternativas existentes en el mejor modelo MNLM, se estimaron dos modelos MLM con componentes aleatorias de error en las funciones de propensión que simulan la correlación entre alternativas.

En segundo lugar, el MNLM no permite estimación de parámetros aleatorios, y a pesar de que el MLM sí lo permite, esto no se realizó debido a que no existe heterogeneidad de los parámetros, ya que el tipo de variables que se usan para estimar los modelos corresponden a factores no inherentes a los conductores. Por lo tanto, esta limitación es propia del MNLM y no para este estudio.

En tercer lugar, el uso de una muestra con razón Caso-Control de 1:5 también genera un problema en los MNLM utilizados, ya que un muestreo del tipo Caso-Control debe ser corregido con el uso de ponderadores en la estimación por máxima verosimilitud.

Teniendo esto en consideración, se procedió a estimar los modelos MLM y en las siguientes secciones se muestran: el cálculo de los ponderadores, la especificación de los modelos, la comparación entre ellos junto a sus resultados, y la interpretación de los parámetros estimados del mejor modelo.

4.2.1 Cálculo de Ponderadores

El cálculo de los ponderadores corresponde a la división entre la probabilidad poblacional de observar un accidente de tipo i y la probabilidad muestral de observar el mismo tipo de accidente. En la Tabla 4-14 se presentan las probabilidades y ponderadores por alternativa.

Tabla 4-14: Probabilidades y Ponderadores.

Tipo de Accidente	Poblacional	Muestral	Ponderadores
NA	0,99984	0,83333	1,19981
I_1p	0,00008	0,08419	0,00100
I_2p	0,00001	0,01572	0,00084
HM_1p	0,00004	0,04500	0,00094
HM_2p	0,00002	0,02175	0,00092

La probabilidad poblacional se obtuvo discretizando la autopista tanto en tiempo como en espacio. En cuanto al horizonte de tiempo (año 2012), este fue agregado en intervalos de 5 minutos, la autopista fue discretizada en celdas de 700 metros. Esta configuración nos da un total de 9.444.384 ensayos. Se observaron solo 1.505 accidentes distribuidos en: 793 del tipo “Sin Heridos y Bloqueo Menor”, 125 del tipo “Sin Heridos y Bloqueo Mayor”, 398 del tipo “Con Heridos y Bloqueo Menor” y 189 del tipo “Con Heridos y Bloqueo Mayor”.

La probabilidad muestral se obtiene por un simple cálculo de la cantidad de veces observadas de un tipo de accidente sobre el total de observaciones en la muestra.

Estos ponderadores se introducen en la base de datos, agregando el ponderador de la alternativa i a cada observación que resultó en un accidente tipo i .

4.2.2 Especificación de los MLM

Esta sección muestra la especificación de los MLM utilizados para representar de mejor manera el fenómeno en estudio. En particular, se explica la manera en que se utilizaron las componentes aleatorias de error para representar la correlación entre las alternativas utilizadas en el mejor MNLM.

Para emplear estas componentes de error, se utilizó como punto de partida el mejor MNLM que corresponde al mostrado en la sección 4.1.7. Estas componentes de error buscan representar correlaciones entre alternativas como lo hace un HLM o un modelo *Cross-Nested Logit* (CNLM). La diferencia entre un HLM y CNLM, es que el HLM no permite correlación entre alternativas del mismo nido, en cambio el CNLM y MLM sí lo permiten. CNLM no permite componentes aleatorias de error y MLM sí, por lo tanto un MLM es más apropiado para esta investigación simulando un *Cross-Nested Probit*.

En este caso, la correlación está dada por DP y DT. En DP se espera la siguiente correlación: “Sin Heridos y Bloqueo Menor” con “Sin Heridos y Bloqueo Mayor”, y “Con Heridos y Bloqueo Menor” con “Con Heridos y Bloqueo Mayor”. En DT la correlación debería ser así: “Sin Heridos y Bloqueo Menor” con “Con Heridos y Bloqueo Menor”, y “Sin Heridos y Bloqueo Mayor” con “Con Heridos y Bloqueo Mayor”. Por lo tanto, la especificación queda formulada como en la Ecuación (4-1).

$$\begin{aligned}
 \text{No Accidente:} \quad & T_{NA} = V_{NA} + \mu_j + \eta_0 + \varepsilon \\
 \text{Sin Heridos y Bloqueo Menor:} \quad & T_{I1p} = V_{I1p} + \mu_I + \eta_{1p} + \varepsilon \\
 \text{Sin Heridos y Bloqueo Mayor:} \quad & T_{I2p} = V_{I2p} + \mu_I + \eta_{2p} + \varepsilon \\
 \text{Con Heridos y Bloqueo Menor:} \quad & T_{HM1p} = V_{HM1p} + \mu_{HM} + \eta_{1p} + \varepsilon \\
 \text{Con Heridos y Bloqueo Mayor:} \quad & T_{HM2p} = V_{HM2p} + \mu_{HM} + \eta_{2p} + \varepsilon
 \end{aligned} \tag{4-1}$$

El término V_i corresponde a la parte determinística de la función de propensión T_i de la alternativa i , mientras que μ y η son las componentes de error aleatorias para representar la correlación según DP y DT, respectivamente. ε es el término de error IID Gumbel.

Por un lado, si consideramos que los términos de error μ y η son normales con media cero y varianzas distintas (tal como lo muestra la Ecuación (4-1)) tenemos un modelo heterocedástico (MLMhe). En cambio, si consideramos que μ es IID Normal($0, \sigma_{DP}^2$) y η IID Normal($0, \sigma_{DT}^2$), tenemos un modelo homocedástico (MLMho).

Estos dos tipos de especificaciones para los MLM surgen a partir de la siguiente interrogante. ¿Es necesariamente la correlación según DP (DT) distinta para alternativas que tienen como resultado distintos niveles de severidad? Es decir, la correlación entre alternativas del tipo “Sin Heridos” (“Bloqueo Menor”) no necesariamente es distinta de la correlación entre alternativas del tipo “Con Heridos” (“Bloqueo Mayor”).

En la sección siguiente se comparan estos modelos y se muestran los resultados.

4.2.3 Comparación y resultados de modelos

Para comparar los modelos, se debe tener en cuenta que deben ser modelos anidados y estimados con el mismo tamaño de muestra para poder aplicar el *Test* de Razón de Verosimilitud. Por lo tanto, para comparar los MLM con el mejor MNLM, se volvió a estimar el MNLM con el uso de ponderadores y la muestra completa, sabiendo que éste es el más restringido de todos.

La reestimación del mejor MNLM con ponderadores (MNLMp) y la muestra completa presenta una log-verosimilitud de -8,809, y los MLMhe y MLMho poseen log-verosimilitudes de -8,805 y -8,809, respectivamente. Esto nos arroja un estadístico LR con valor 0,008 si comparamos MNLMp con MLMhe, y 0 si comparamos MNLMp con MLMho. Ambos LR son menores que $\chi_{6;95\%}^2=12,6$ y $\chi_{2;95\%}^2=5,9$, y por lo tanto no se puede decir que el modelo más restringido es erróneo.

La comparación entre MLMhe y MLMho se realizó sabiendo que el modelo homocedástico es restringido del heterocedástico. La prueba arrojó un valor de 0,008 para el estadístico LR, y este no es mayor que el valor de tabla de una distribución $\chi_{4;95\%}^2 = 9,5$. Por lo tanto, no se puede asegurar que el MLMho es erróneo.

Hasta este punto, ninguno de los modelos es estadísticamente erróneo según el análisis basado en el Test de Razón de Verosimilitud. Por lo tanto, en términos de log-verosimilitud el MLMhe es mejor. En la Tabla 4-15, Tabla 4-16 y Tabla 4-17, se

muestran los resultados de estos tres modelos, y luego se analizan los parámetros de las estructuras de error.

El modelo MNLMp no posee estructuras de error, por lo que es descartado. En cambio, los modelos MLMhe y MLMho sí consideran correlación entre las alternativas por medio de las componentes de error y se debe verificar qué estructura es la correcta. Para ello, se deben observar los parámetros asociados a los errores en los modelos MLMhe y MLMho, presentes en la Tabla 4-16 y Tabla 4-17, respectivamente.

La interrogante en cuanto a los distintos tipos de especificación según correlación se responde al mirar el nivel de significancia de las desviaciones estándar de las componentes de error. El modelo MLMho no presenta ninguno de los parámetros significativos en cuanto a este aspecto. Sin embargo, el MLMhe presenta parámetros significativos al menos al 65% de confianza, y por lo tanto, se justifica el uso de estructuras de error para correlacionar alternativas. Según el modelador, un nivel de 65% de confianza es aceptable.

Considerando los aspectos a favor y en contra para cada uno de los modelos, se cree que el mejor modelo debería ser el MLMhe, ya que es el modelo que entrega valor y representa bien el fenómeno.

Por otro lado, se puede decir que los datos indican la existencia de factores no observables que correlacionen las alternativas, y por lo tanto, considerar una estructura de error para la correlación entre alternativas es aceptable. Además, si se comparan los valores de los parámetros en cada uno de los modelos, no son muy distantes entre ellos.

Tabla 4-15: Resultados de Estimación con muestra completa – MNLMp.

Parámetro	Valor	Valor-p	Riesgo Relativo
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,819	0,00	2,268
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,803	0,00	0,448
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	0,680	0,00	1,974
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	0,199	0,45	1,220
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,052	0,00	1,053
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,078	0,00	1,081
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,051	0,00	1,053
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,110	0,00	1,116
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,009	0,03	1,009
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,076	0,00	1,079
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,006	0,23	1,006
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,015	0,01	1,015
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,170	0,07	0,844
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,674	0,00	0,510
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,790	0,00	16,250
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	1,260	0,00	3,519
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	2,140	0,00	8,483
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	2,670	0,00	14,412
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,850	0,00	6,355
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-1,390	0,01	0,249
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,702	0,00	2,016
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,780	0,00	5,924
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,610	0,00	4,998
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	1,940	0,00	6,952
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,190	0,00	3,284
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,0018	0,00	1,0018
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	0,942	0,06	2,565
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,682	0,00	0,506
SEMANA (Con Heridos y Bloqueo Mayor)	-0,153	0,32	0,858
CONSTANTE (Con Heridos y Bloqueo Menor)	-11,200	0,00	--
CONSTANTE (Con Heridos y Bloqueo Mayor)	-12,500	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Menor)	-10,800	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-14,400	0,00	--
CONSTANTE (No Accidente)	0	--	--
Log-verosimilitud	-8,805		

Tabla 4-16: Resultados de Estimación con muestra completa – MLM Heterocedástico.

Parámetro	Valor	Valor-p	Riesgo Relativo
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,796	0,01	2,210
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,812	0,02	0,444
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	0,687	0,03	1,987
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	0,212	0,77	1,236
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,052	0,00	1,054
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,078	0,00	1,082
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,051	0,00	1,052
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,110	0,00	1,116
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,009	0,33	1,009 (0,999)*
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,076	0,00	1,079
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,007	0,58	1,007 (0,999)*
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,017	0,27	1,017 (0,999)*
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,195	0,20	0,823
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,646	0,01	0,525
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,790	0,00	16,190
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	1,260	0,01	3,507
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	2,170	0,00	8,571
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	2,650	0,00	13,844
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,850	0,00	6,344
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-1,390	0,12	0,249
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,741	0,03	2,089
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,790	0,00	5,972
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,600	0,01	4,939
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	1,980	0,00	7,134
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,240	0,00	3,406
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,0018	0,00	1,0018
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	0,958	0,18	2,606
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,677	0,01	0,508
SEMANA (Con Heridos y Bloqueo Mayor)	-0,196	0,46	0,822
μ_j (Desviación Estándar No Accidente)	0,707	0,18	--
μ_I (Desviación Estándar Sin Heridos)	0,332	0,35	--
μ_{HM} (Desviación Estándar Con Heridos)	0,566	0,07	--
η_0 (Desviación Estándar No Accidente)	1,930	0,00	--
η_{1p} (Desviación Estándar Bloqueo Menor)	0,754	0,18	--
η_{2p} (Desviación Estándar Bloqueo Mayor)	0,401	0,22	--
CONSTANTE (Con Heridos y Bloqueo Menor)	-13,400	0,00	--
CONSTANTE (Con Heridos y Bloqueo Mayor)	-14,900	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Menor)	-11,500	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-15,200	0,00	--
CONSTANTE (No Accidente)	0	--	--
Log-verosimilitud	-8,805		

* Valor de Riesgo relativo cuando el parámetro no está en el modelo.

Tabla 4-17: Resultados de Estimación con muestra completa – MLM Homocedástico.

Parámetro	Valor	Valor-p	Riesgo Relativo
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,820	0,01	2,270
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,805	0,02	0,447
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	0,683	0,03	1,980
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	0,201	0,78	1,223
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,052	0,00	1,054
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,078	0,00	1,081
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,052	0,00	1,053
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,110	0,00	1,116
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,009	0,33	1,009
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,077	0,00	1,079
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,006	0,60	1,006
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,015	0,32	1,015
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,172	0,26	0,842
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,672	0,01	0,511
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,790	0,00	16,258
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	1,260	0,01	3,520
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	2,140	0,00	8,487
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	2,670	0,00	14,419
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,850	0,00	6,356
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-1,390	0,12	0,249
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,700	0,04	2,013
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,780	0,00	5,925
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,600	0,01	4,949
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	1,940	0,00	6,954
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,190	0,00	3,285
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,0018	0,00	1,0018
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	0,948	0,18	2,580
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,682	0,01	0,506
SEMANA (Con Heridos y Bloqueo Mayor)	-0,158	0,56	0,854
μ (Desviación Estándar Daños Personales)	0,707	0,61	--
η (Desviación Estándar Disrupción del Tráfico)	0,332	0,81	--
CONSTANTE (Con Heridos y Bloqueo Menor)	-11,500	0,00	--
CONSTANTE (Con Heridos y Bloqueo Mayor)	-12,800	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Menor)	-11,100	0,00	--
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-14,700	0,00	--
CONSTANTE (No Accidente)	0	--	--
Log-verosimilitud	-8,809		

4.3 Análisis e Interpretación de Parámetros de Mejor Modelo

Por las razones mencionadas, el mejor modelo corresponde al MLMhe presentado en la Tabla 4-16. Por lo tanto, el análisis e interpretación de los parámetros se realiza sobre este, utilizando el riesgo relativo.

Cada análisis con el uso del riesgo relativo de un parámetro, puede indicar un aumento (disminución) de probabilidades de ocurrencia de cualquiera de los tipos de accidentes, considerando que el resto de los parámetros permanecen constantes.

Además, se recuerda que este modelo estima probabilidades de ocurrencia de accidentes menores que los MNLM sin ponderadores, y por lo tanto aumentos muy grandes no significa que la probabilidad aumente demasiado, ya que se compara con otra probabilidad que ya es muy pequeña.

4.3.1 Significancia, Signo y Magnitud de los Parámetros

En general, se muestran resultados que consideran los mismo parámetros que el mejor MNLM sin ponderadores, pero en éste caso hay algunos de ellos que no son significativos dentro de lo aceptado (80% de confianza). No obstante, debido a que no existen correlaciones altas entre parámetros asociados a las variables, se dejan estas variables en el modelo, ya que sacarlas no afecta el valor estimado de los parámetros, y por lo tanto los riesgos relativos permanecen constantes.

Los signos y magnitudes se comportan de acuerdo a lo esperado, tal como en los MNLM.

4.3.2 Parámetros asociados a interacciones

Los resultados del MLMhe indican que la presencia del Estado “Final de Cola” en un tramo con pendiente de bajada, la probabilidad de ocurrencia de un accidente “Con Heridos y Bloqueo Mayor” es 2,21 veces mayor que en un tramo sin pendiente y/o el Estado “Final de Cola”. Bham et al. (2012) mencionan que la presencia de gradiente aumenta la severidad de un choque 2,25 veces en comparación con secciones niveladas.

Por su parte, manteniendo una pendiente de bajada, pero con una condición de tráfico congestionada, la probabilidad de ocurrencia de un accidente “Sin Heridos y

Bloqueo Menor” disminuye a 0,44 veces la probabilidad cuando no existen pendiente y/o el Estado “Congestión”. En cambio, si la condición correspondiese a la del Estado “Cuello de Botella”, la probabilidad de ocurrencia de un accidente de tipo “Sin Heridos y Bloqueo Menor” se vuelve 1,99 veces mayor en comparación a una situación sin pendiente y/o el Estado “Cuello de Botella”.

Con respecto al clima, en la literatura algunos estudios mencionan que indicadores de estación del año son significativos, implicando que las condiciones climáticas pueden afectar la severidad de un accidente (Jung et al., 2010). Por un lado, Yang et al. (2011) dicen que una buena iluminación y condiciones favorables del clima disminuyen la probabilidad de ocurrencia de un accidente severo, mientras que Eluru & Bhat (2007) mencionan que tener condiciones pobres de iluminación y adversas del clima indican lesiones poco severas, puesto que los conductores están más alerta.

Por el lado de los factores geométricos, estudios previos, en el marco de modelos de elección discreta, mencionan que la presencia de curvas en una sección de la vía aumenta la probabilidad de tener accidentes más perjudiciales capaces de producir lesiones (Abdel-Aty, 2003; Khattak et al., 2002). Bham et al. (2012), específicamente, menciona que la severidad aumenta 2,29 veces cuando la ocurrencia de un choque fue en presencia de una curva en comparación con una sección recta.

En este estudio se consideró un efecto conjunto que captura factores climáticos y geométricos por medio de la variable “LLUVIAxCURVA”. Este efecto corresponde al que produce la lluvia en un tramo que presenta curvas. Su riesgo relativo promedio indica que cuando esta variable es igual a 1, la probabilidad de ocurrencia de la alternativa “Sin Heridos y Bloqueo Menor” es 2,6 veces mayor en comparación a una situación sin lluvia y/o el tramo no presentase curvas.

4.3.3 Parámetros asociados a Desviaciones Estándar de Velocidad

Con respecto a los parámetros asociados a las variables que miden las desviaciones estándar de las velocidades tanto aguas arriba como aguas abajo, éstos aumentan la propensión de ocurrencia de cualquiera de los tipos de accidentes. Es decir, a medida que las desviaciones aumentan también lo hacen las probabilidades de accidente.

Para los parámetros asociados a la estación 1, se observa que un aumento en una unidad en las desviaciones estándar de aguas arriba, las probabilidades de los accidentes “Sin Heridos y Bloqueo Menor”, “Sin Heridos y Bloqueo Mayor”, “Con Heridos y Bloqueo Menor” y “Con Heridos y Bloqueo Mayor” aumentan en 5,4%, 8,2%, 5,2% y 11,6%, respectivamente. Por otra parte, para el mismo aumento, pero en las desviaciones estándar de la estación de aguas abajo (estación 3), las probabilidades de ocurrencia de accidentes “Sin Heridos y Bloqueo Menor”, “Sin Heridos y Bloqueo Mayor”, “Con Heridos y Bloqueo Menor” y “Con Heridos y Bloqueo Mayor” aumentan en 0,9%, 7,9%, 0,7% y 1,7%, respectivamente.

La variabilidad de la velocidad en la estación 1 indica, claramente, una mayor relevancia en comparación con la desviación estándar de la velocidad en la estación 3. Esto se debe a que las variaciones de velocidad en lugares más cercanos a los vehículos afectan de mayor manera, siendo siempre la estación 1 la más cercana. En general, estos resultados pueden indicar que los conductores ajustan sus velocidades de manera más frecuente a medida que la variabilidad de la velocidad aumenta, lo cual induce en ellos más posibilidad de error en mantener la distancia de seguridad entre vehículos (Lee et al., 2002).

4.3.4 Parámetros asociados a Estados de Tráfico

Al igual que en los modelos *Logit* Multinomial sin ponderadores, los parámetros asociados a los Estados de Tráfico, vuelven a indicar, en general, que los estados de transición son los más riesgosos con respecto al estado base. Los riesgos relativos de los Estados de Tráfico indican las siguientes variaciones en las probabilidades:

- Para un accidente “Sin Heridos y Bloqueo Menor” con los Estados “Final de Cola”, “Congestión” o “Cuello de Botella” presentes, las probabilidades son 16,2, 6,3 y 6 veces mayor con respecto al Estado “Flujo Libre” (base), respectivamente.
- Para un accidente “Sin Heridos y Bloqueo Mayor” con los Estados “Final de Cola”, “Congestión” o “Cuello de Botella” iguales a 1, las probabilidades son 3,5, 0,25 y 4,9 veces la probabilidad obtenida si el estado base estuviese activo.

- Para un accidente “Con Heridos y Bloqueo Menor” con los Estados “Final de Cola”, “Congestión” o “Cuello de Botella” presentes, las probabilidades son 8,5, 2,1 y 7,1 veces mayor con respecto al Estado “Flujo Libre”, respectivamente.
- Para un accidente “Con Heridos y Bloqueo Mayor” con los Estados “Final de Cola” o “Cuello de Botella” activados (iguales a 1), las probabilidades son 13,8 y 3,4 veces mayores que las probabilidades cuando los Estados “Flujo Libre” o “Congestión” son iguales a 1, respectivamente. En este caso, el Estado “Congestión” no pareció comportarse significativamente diferente del Estado “Flujo Libre”.

De lo anterior, se aprecia que la probabilidad de ocurrencia más afectada es la que corresponde al accidente “Sin Heridos y Bloqueo Menor”. Esto se debe a que este tipo de accidente es el más frecuente, y es lógico pensar que su probabilidad aumente en mayor magnitud en comparación a las demás alternativas.

4.3.5 Parámetros asociados a Variables Geométricas

El parámetro asociado a la variable ENTRADA indica una disminución en la propensión de ocurrencia de un accidente de tipo “Con Heridos y Bloqueo Mayor”. El riesgo relativo indica que esta probabilidad disminuye en un 47,5% cuando un tramo posee una pista de entrada a la autopista. Como se explicó en análisis anteriores, el simple hecho de que los conductores en la vía se percaten de la presencia de una pista de entrada, hace que se concentren más y conduzcan de manera más precavida.

En el caso de la variable “DISTANCIA ENTRE ESTACIONES 1 Y 3”, su parámetro estimado indica un aumento en la propensión de ocurrencia de un accidente “Sin Heridos y Bloqueo Mayor”. En detalle, cuando la variable “DISTANCIA ENTRE ESTACIONES 1 Y 3” aumenta en una unidad, la probabilidad de ocurrencia de un accidente “Sin Heridos y Bloqueo Mayor” aumenta en 0,2%.

Por otro lado, si seguimos analizando la alternativa “Sin Heridos y Bloqueo Mayor”, el parámetro asociado a la variable “PENDIENTE EN SUBIDA” dicta una disminución en el riesgo de ocurrencia de ésta. La presencia de una pendiente en subida

disminuye la probabilidad de un accidente “Sin Heridos y Bloqueo Mayor” en un 49,2%, si consideramos *certeris paribus* para el resto de las variables.

4.3.6 Resto de los Parámetros

El parámetro estimado, específico de la alternativa “Con Heridos y Bloqueo Menor”, asociado a “DURANTE EL DÍA EN DÍA DE SEMANA”, indica que este tipo de accidentes reduce su riesgo cuando se está en un día laboral y es de día. De hecho, disminuye su probabilidad de ocurrencia en un 17,7%. Esto es esperable, puesto que de noche y/o fin de semana, se tiende a circular a velocidades mayores por tener menos flujo en la vía (Eluru & Bhat, 2007), lo cual está directamente relacionado con la severidad según la dimensión de daños personales.

Por último, el parámetro asociado a la variable SEMANA se mostró ser estadísticamente significativo en la alternativa “Con Heridos y Bloqueo Mayor”, indicando una disminución en la propensión de ésta. Si SEMANA fuese igual a 1, la probabilidad de ocurrencia de un accidente “Con Heridos y Bloqueo Mayor” disminuye en un 17,8%. Por lo tanto, se cree que en días laborales, la velocidad es menor que en los fines de semana, generando menos posibilidad de severidad, tanto en daños personales como en interrupción del flujo vehicular. Se esperaría que esta variable también afectase a la alternativa “Con Heridos y Bloqueo Menor”, pero ya es afectada por el parámetro asociado a DIA_SEM, que posee una fuerte correlación la variable SEMANA.

5 PROPUESTAS BASADAS EN LOS RESULTADOS

En este capítulo se exponen propuestas basadas en los resultados que buscan prevenir o asistir de mejor manera la ocurrencia de accidentes en autopistas, específicamente en la Autopista Central de Santiago de Chile. Estas propuestas corresponden a sugerencias que podrían eventualmente ser evaluadas para su posible implementación.

En esta investigación, las variables consideradas influyentes en la ocurrencia de accidentes en autopistas corresponden a factores observables y no inherentes a los conductores. Así, las propuestas se derivan a partir de estos factores observables y controlables por los operadores de la autopista, que en la mayoría de los casos, corresponden a variables de tráfico. Por otro lado, tanto factores climáticos como geométricos, si bien son observables, no son controlables por los operadores en el corto plazo.

Las propuestas se pueden clasificar en dos tipos: Estrategias de Asistencia y Estrategias de Prevención; las cuales no son necesariamente excluyentes entre ellas.

A diferencia de operadores o entidades gubernamentales que actúan sobre redes viales urbanas con flujo interrumpido, los operadores de autopistas no tienen mucha posibilidad de intervención, ya que las autopistas no poseen intersecciones semaforizadas ni prioritarias, y por lo tanto la señalética existente es mucho menor. Por lo tanto, el universo de estrategias de asistencia o prevención es más reducido.

Para llevar a cabo estas propuestas, se recomienda, fuertemente, el monitoreo de los tramos existentes en la Autopista Central. Estos tramos corresponden a los pares de espiras magnéticas consecutivas espacialmente, es decir, los tramos utilizados en esta investigación para definir las observaciones en la bases de datos utilizadas para estimar los modelos (ver sección 3.2).

A continuación se presentan las propuestas de Estrategias de Asistencia, y luego aquellas propuestas de Estrategias de Prevención.

5.1 Estrategias de Asistencia

Las estrategias de asistencia son aquellas que apuntan a brindar una pronta asistencia a vehículos que sufren accidentes de algún tipo. Esta investigación entrega un modelo que permite el cálculo de probabilidades de ocurrencia de accidentes según su severidad tanto en daños personales como en disrupción en el tráfico. Por lo tanto, la asistencia no es netamente del tipo médica, sino que también, es de carácter técnico.

En este aspecto de asistencia, las propuestas apuntan principalmente a la localización de los recursos (ambulancias, grúas, entre otros) que la autopista posee para asistir accidentes. Por lo tanto, el monitoreo de los tramos, o más bien, el monitoreo de las variables consideradas influyentes en esta investigación de un cierto tramo, podrían dar un indicio de cuando un tramo se vuelve potencialmente un lugar donde vaya a ocurrir un accidente. Si este monitoreo se realiza en un periodo de tiempo lo suficientemente largo, se puede estudiar la posibilidad de localizar de mejor forma los recursos utilizados para asistencia en seguridad vial a lo largo de la Autopista Central, y minimizar el tiempo de llegada al lugar de un incidente. Esto puede minimizar la probabilidad de una fatalidad a causa de una herida grave, y también las demoras que se pueden generar.

Una limitación que este tipo de monitoreo posee, es que se asume que los lugares riesgosos o *hotspots* permanecen estáticos, lo cual no es necesariamente correcto, ya que la demanda vehicular puede ser dinámica, y por lo tanto, el tráfico también lo sería. Esto produce que los *hotspots* cambien con el tiempo. Dada esta limitación, el monitoreo podría realizarse con alta frecuencia (cortos intervalos de tiempo). Por ejemplo, este intervalo de tiempo podría ser de cinco minutos, ya que es el mismo largo del intervalo de tiempo que se utilizó en esta investigación para definir las variables de tráfico.

No es sensato proponer que la autopista relocalice sus recursos de asistencia cada cinco minutos, pero un monitoreo de este tipo permite evaluar el reajuste de los vehículos con una frecuencia razonable, ya sea diariamente como semanalmente, según el operador estime conveniente.

Si el monitoreo se realiza en intervalos cortos de tiempo, se puede identificar, casi en tiempo real, el comportamiento de los tramos en la autopista y definir un umbral de probabilidad sobre el cual los operadores pueden, dentro de lo posible, reaccionar mejor para asistir un accidente. Por ejemplo, la identificación de lugares riesgosos con este monitoreo permite mejores decisiones por parte de los operadores en cuanto a asistencia en la vía, indicando de qué lugar es mejor que salgan los vehículos de asistencia, o bien, pueden identificar lugares congestionados y tomar medidas al respecto.

5.2 Estrategias de Prevención

Debido a la limitada posibilidad de intervención por parte de los operadores, la gran mayoría de las estrategias de prevención que se proponen en esta sección corresponden a formas de entregar información a los usuarios.

Al igual que en las estrategias de asistencia, las estrategias de prevención también requieren del seguimiento de los indicadores de probabilidad de cada tramo, ya que con éste se pueden implementar sistemas inteligentes para estabilizar la velocidad (es decir, mantener la velocidad constante), o suavizar la transición entre velocidades altas y bajas, y disminuir el riesgo de accidentes. Estas medidas apuntan a disminuir, dentro de lo posible, el riesgo inmerso en Estados de Tráfico de transición (Estados 1 o 4). Estos estados presentan el mayor riesgo relativo, y por ende el mayor aumento en las probabilidades de ocurrencia de accidentes de cualquier tipo, ya sea tanto en DP como en DT.

Dado que los Estados de Tráfico, resultan ser los factores mayormente influyentes en la ocurrencia de accidentes, se debe tener en cuenta para cada propuesta las distintas formas de identificación de los estados en cada tramo. Para ello se pueden usar enfoques de estimación de los Estados de Tráfico o simplemente identificar los observados en cada tramo cada cinco minutos.

También, según la Figura 5-1, se debe tener en cuenta las posibles transiciones entre Estados de Tráfico. Los operadores deberán tener presente el Estado de Tráfico actual, así como también el futuro estado que se podría observar. Se debe destacar que

los Estados de Transición pueden pasar desde uno al otro sin necesariamente tener que pasar por un Estado de Flujo Libre o Congestión.

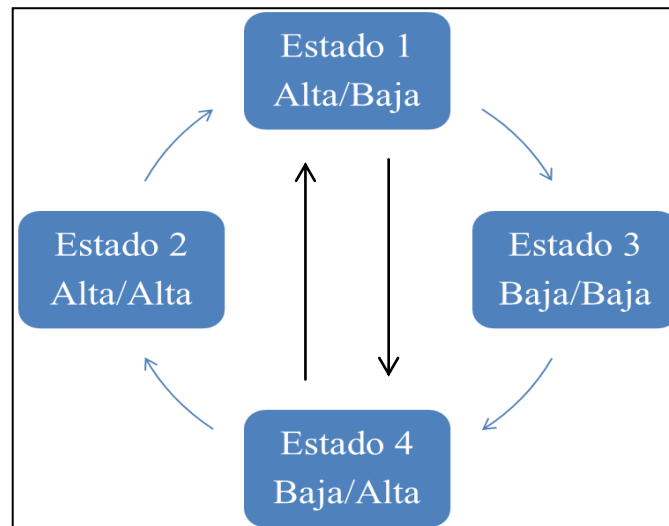


Figura 5-1: Cambios en los Estados de Tráfico.

En términos generales, estas estrategias pueden ser estudiadas para evaluar su posible implementación en el corto o largo plazo, dependiendo de la medida en específico. Dentro de las propuestas que se presentan, se consideran: el control de acceso (o *ramp metering*), límites de velocidad variable (*Variable Speed Limits* o VSL), paneles de mensaje variable y estrategias de diseño.

5.2.1 *Ramp Metering*

Los sistemas de *Ramp Metering* corresponden al uso de dispositivos que dosifican el flujo de entrada a autopistas de acuerdo con las condiciones actuales del tráfico dentro de las autopistas. Estos dispositivos son semáforos en las rampas de entradas que regulan la tasa de entrada de vehículos a la autopista. Los sistemas de *Ramp Metering* tienen como objetivo disminuir la congestión y mejorar la seguridad vial.

El monitoreo de los tramos permite saber los Estados de Tráfico presentes en tramos que posean pistas de entradas, y por lo tanto, el uso de esta información sirve como variable de entrada a estos sistemas, además del flujo actual en esas zonas. Por lo tanto, las sugerencias en este aspecto indican que si en un instante de tiempo, el tramo

con pista de entrada a la autopista presenta un estado de transición, se debe restringir el acceso de vehículos a la autopista. Además, se debe restringir el acceso de vehículos para estados congestionados (Estado 3), ya que, además de ser riesgoso, *ramp metering* busca disminuir la congestión.

Si bien en el Capítulo 4 se mencionó que la variable ENTRADA disminuía la probabilidad de accidentes, esta disminución no se justifica si sabemos que los Estados de Tráfico de transición considerablemente aumentan de mayor manera las probabilidades.

Debido a que generalmente, se deben asignar los recursos de manera óptima, por ser estos escasos, se puede usar el monitoreo de tramos con pistas de entrada por un largo tiempo, para reconocer aquellas zonas más riesgosa que necesiten más del uso de este tipo de sistema. Sin embargo, nuevamente nos encontramos con la limitación de asumir que las zonas riesgosas permanecen fijas en el tiempo.

5.2.2 Límites de Velocidad Variable (*Variable Speed Limits o VSL*)

Algunos autores (Kononov, Durso, et al., 2012; Kononov, Hersey, et al., 2012) recomiendan el uso de límites de velocidad variable (*Variable Speed Limits o VSL*), es decir, distintos niveles de velocidad máxima a lo largo de la autopista que cambian con el tiempo. Lee et al. (2006) desarrollaron un modelo que evalúa los beneficios de usar VSL en términos de seguridad vial y demuestran que puede reducir el potencial de accidente desde un 5 hasta un 17%.

Los Estados de Tráfico no se pueden cambiar a voluntad por ser son condiciones de operación de la autopista dadas por la demanda vehicular. Sin embargo, se pueden incentivar ciertos comportamientos que propicien ciertos estados. La finalidad del uso de límites de velocidad variable precisamente busca incentivar comportamientos que induzcan estados de flujo libre y velocidades constantes. En casos en que esto no se logre, se obtendrán estados congestionados y de transición, donde el objetivo es que dichas transiciones de la velocidad se hagan de manera poco brusca. Además, esta estrategia podría disminuir la variabilidad de la velocidad que también aumenta el riesgo de los accidentes.

En los siguientes párrafos se presentan sugerencias de cómo se podría implementar VSL basado en el monitoreo de los Estados de Tráfico y cómo puede ayudar.

Con el equipamiento disponible de paneles de mensaje variable en la autopista, los tramos definidos por espiras espacialmente consecutivas se podrían agregar en tramos entre paneles, y presentar un mensaje en aquellos paneles que sugieran velocidades máximas de circulación, cuando dentro de estos nuevos tramos se detecten estados de congestión, o bien, de transición, que son los que aumentan en la mayoría de los casos las probabilidades de ocurrencia de accidentes.

5.2.3 Paneles de Mensaje Variable y Mensajes Directos al Usuario

Los Paneles de Mensaje Variable (PMV) es un equipamiento que las autopistas poseen para dar información a los usuarios por medio de mensajes directos que pueden cambiar según el operador estime conveniente.

Con los resultados obtenidos se pueden realizar dos cosas: determinar lugares donde es conveniente instalar los PMV y diseñar mensajes que prevengan la ocurrencia de accidentes.

En primer lugar, la determinación de los mejores lugares para instalar los PMV requiere del monitoreo de los tramos en un periodo largo de tiempo, por lo tanto se considera como una estrategia de largo plazo, y nuevamente se asume que los *hotspots* permanecen fijos en el tiempo.

En segundo lugar, el diseño de mensajes directos a los usuarios tiene como objetivo transformar zonas riesgosas en zonas más seguras (o menos riesgosas), por medio de los PMV. El diseño de mensajes no es una tarea trivial, y puede ser una tarea que se debe llevar a cabo en conjunto con expertos en psicología para aproximarse a los usuarios de la mejor manera.

A continuación, se detallan cuáles deberían ser los objetivos de estos mensajes, que la autopista debería proveer a los usuarios, ante Estados de Tráfico riesgosos:

- Estado 1: los vehículos están entrando a una zona congestionada, y por lo tanto, el mensaje debería advertir a los usuarios de la presencia de congestión, con el fin de evitarla (re-ruteo), o bien, persuadir mayor concentración en ellos.

También, el mensaje podría indicar cambios en la conducción, ya sea modificando la velocidad o la distancia entre vehículos (espaciamiento).

- Estado 4: los vehículos están saliendo de una zona congestionada, por lo que se debería exponer un mensaje que no incite aceleraciones descontroladas que cambien bruscamente la velocidad, ya que este cambio es el que genera riesgo de accidentes de cualquier tipo.
- Estado 3: los vehículos están en una zona congestionada (cola). Si bien este estado es menos riesgoso que los estados 1 y 4, aumenta en mayor magnitud la probabilidad de accidentes sin DP y bajo bloqueo de pistas, debidos al *stop and go*. Por lo tanto, un mensaje que permita aumentar la concentración de los conductores, o bien, incitarlos a mantener un espaciamiento prudente, podría prevenir este tipo de accidentes.

Además, el uso de PMV para dar información a los usuarios, también, podría exponer mensajes basados en los demás factores influyentes, como lo son la geometría y el clima. Por ejemplo, un mensaje que indique concentración y disminución de velocidad por la presencia de lluvia en tramos con curvas.

5.2.4 Estrategias de Diseño

Las estrategias de diseño se cree que pueden ser implementadas en el largo plazo, ya que requieren de estudios más exhaustivos sobre los resultados obtenidos en esta investigación.

En primera instancia, las estrategias de diseño pueden transformarse en herramientas que los ingenieros y diseñadores deberían tener en cuenta al momento de diseñar un proyecto de infraestructura vial, como lo son las autopistas urbanas.

En segundo lugar, el estudio de este tipo de estrategias puede relacionarse con el tipo de demarcaciones existentes en las autopistas para prevenir la ocurrencia de accidentes. Por ejemplo, las demarcaciones en forma de *zigzag* en los costados de una pista incitan la sensación de estrechamiento de la pista que inducen velocidades menores. También, el diseño del entorno de la autopista, es decir, el diseño de los alrededores de ella, podría tenerse en cuenta para no generar distracciones en los

usuarios. Estas distracciones podrían eventualmente causar Estados de Tráfico riesgosos que llevan a accidentes.

6 CONCLUSIONES Y RECOMENDACIONES

6.1 Conclusiones Generales

El objetivo general de esta investigación fue formular un modelo que permita estimar, basado en Estados de Tráfico observados y otros factores fácilmente observables y no inherentes a los conductores, probabilidades de ocurrencia de distintos tipos de accidente diferenciados en función de su severidad según daños personales y interrupción en el tráfico, considerando, también, la alternativa que no ocurra un accidente. Para esto se estimaron modelos de elección discreta con muestras del tipo Caso-Control.

Se estimaron dos tipos de modelos: *Logit Multinomial* (MNLM) y *Logit Mixto* (MLM) con componentes de error. El uso de los primeros fue netamente para verificar la influencia de las variables en distintas especificaciones, y comprobar que un modelo con especificación de alternativas definidas según daños personales y interrupción en el tráfico, de manera conjunta, es un modelo razonable. También se verificó el tipo de muestreo Caso-Control considerando cinco Controles por Caso (es decir, usar una razón Caso-Control de 1:5).

La utilización de la metodología Caso-Control con MNLM tiene netamente un fin exploratorio, de forma tal de verificar que un modelo con una especificación compleja es capaz de entregar resultados lo suficientemente satisfactorios como para asegurar que la utilización de dicho tipo de especificación apunta en la dirección correcta.

Dado que el modelo anterior apunta en la dirección correcta, se procedió a estimar modelos *Logit Mixto* con componentes de error para solucionar las limitaciones presentes al utilizar un MNLM con este tipo de muestreo y datos, simulando un modelo *Cross-Nested Probit*. Estas limitaciones son: el uso de un muestreo Caso-Control puede sobrestimar las probabilidades, y no considerar correlación entre alternativas podría no representar de mejor manera el fenómeno estudiado. Por lo tanto, se corrigieron estas limitaciones con el uso de ponderadores, y especificaciones de los modelos utilizando componentes de error para representar las correlaciones entre las alternativas y verificar que el uso de estas estructuras de error son correctas.

Se estimaron tres modelos: un MNLM con ponderadores (MNLMP) y dos MLM con componentes de error (MLMhe y MLMho). La comparación entre ellos según un test de Razón de Verosimilitud indicó que todos los modelos son estadísticamente correctos. Por otro lado, la comparación de los resultados entre los modelos MLMhe y MLMho indica que el modelo que representa bien el fenómeno de correlación entre alternativas es el MLMhe, con una significancia de sus parámetros al 65% de confianza. Por lo tanto, se consideró que hay evidencia suficiente para decir que existen factores no observables que correlacionen las alternativas, y por lo tanto, se deja como mejor modelo el MLMhe.

Con MLMhe se puede establecer la verificación de la hipótesis de esta investigación, ya que según los resultados obtenidos, el modelo arroja que los parámetros asociados a los Estados de Tráfico son estadísticamente significativos distintos de cero. Es decir, son influyentes en la ocurrencia de los tipos de accidentes mencionados.

A continuación revisaremos las conclusiones asociadas con los resultados y variables más relevantes de la estimación del modelo. Luego, se mencionan algunas recomendaciones y direcciones para trabajo futuro.

6.2 Relevancia de los Estados de Tráfico

La relevancia de los Estados de Tráfico en la ocurrencia de accidentes, se midió a partir del riesgo relativo promedio calculado para cada uno de los parámetros asociados a estas variables. En comparación con el estado de “Flujo Libre” (Estado 2 o base), los resultados del modelo indican que la gran mayoría de los parámetros asociados a los Estados de Tráfico aumentan la propensión de ocurrencia de los tipos de accidentes definidos. El estado de Flujo Libre fue considerado como base debido a que fue considerado *a priori* como el menos riesgoso, por tener velocidades altas y constantes a lo largo de un tramo específico, y por ser el estado presente en la gran mayoría de las situaciones de no accidente.

Se encontró que los estados de transición son los que presentan mayor riesgo de ocurrencia de accidentes. Es decir, los estados donde los vehículos disminuyen sus

velocidades altas a bajas para incorporarse al “final de una cola” (Estado 1) ocasionada por congestión y la aumentan para liberarse de una como un “Cuellos de Botella” (Estado 4), las probabilidades de ocurrencia de los accidentes definidos aumentan de mayor manera en comparación con el resto de los Estados de Tráfico.

Por ejemplo, la presencia de un estado “Final de Cola” indica, en promedio, un aumento de aproximadamente 11 veces la probabilidad de ocurrencia de los tipos de accidentes definidos. En cambio, para un estado “Cuello de Botella” este aumento vuelve las probabilidades cerca de 5 veces mayor en comparación con un estado de “Flujo Libre”. Para un estado en “Congestión” (Estado 3) se concluyó también un aumento en las probabilidades, pero existe un caso donde éste muestra una disminución en la probabilidad de ocurrencia de un accidente que resulta solo daños materiales y bloqueo de 2 o más pistas (“Sin Heridos y Bloqueo Mayor”). Lo cual es sensato si se piensa que a bajas velocidades es menos probable la ocurrencia de un accidente que bloquee más de una pista.

Dado que el Estado de Tráfico “Final de Cola” es el más riesgoso, se indagó más en su influencia en las probabilidades de ocurrencia de accidentes. Se concluyó que las probabilidades más afectas son aquellas relacionadas, primero, con accidentes que no provocan heridos ni muertos y el bloqueo de pistas es menor (“Sin Heridos y Bloqueo Menor”), seguidos de aquellos con al menos un herido o víctima fatal y bloqueo mayor de pistas (“Con Heridos y Bloqueo Mayor”). Esto se debe al riesgo que presenta para los vehículos una desaceleración obligada por congestión aguas abajo, para disminuir sus velocidades desde altas a bajas.

Con respecto al estado “Congestión”, se concluyó que la probabilidad más afectada es la menos severa tanto en términos de daños personales como en interrupción del tráfico (“Sin Heridos y Bloqueo Menor”), debido a que una condición de tráfico en donde los vehículos circulan a velocidades bajas es poco probable que los accidentes provoquen heridos o muertos y bloqueos de más de una pista.

La aceleración que conlleva un estado “Cuello de botella” que los vehículos realizan para liberarse de una cola, genera la segunda condición de operación más riesgosa, y debido a esto las probabilidades de ocurrencia de accidentes que aumentan en

mayor magnitud son aquellas donde los accidentes producen al menos un herido o muerto y bloqueo de pista menor (“Con Heridos y Bloqueo Menor”), y otra que genera el mismo tipo de bloqueo con daños solo materiales (“Sin Heridos y Bloqueo Menor”).

A continuación se presentan las conclusiones relacionadas con el resto de las variables de control utilizadas en este estudio.

6.3 Relevancia de otras Variables de Control

Con respecto al resto de los factores influyentes, se encontró que algunos pueden contribuir a la ocurrencia de accidentes, y otros pueden disminuir el riesgo. Aquellos que lo disminuyen, no lo hacen de una manera tal que contrarreste el aumento provocado por los Estados de Tráfico.

En lo específico:

- La variabilidad de la velocidad en la estación de aguas arriba (estación 1) indica, claramente, una mayor relevancia en comparación con la desviación estándar de la velocidad en la estación aguas abajo (estación 3). Esto se debe a que las variaciones de velocidad en lugares más cercanos a los vehículos afectan de mayor manera, siendo siempre la estación 1 la más cercana. En general, estos resultados pueden indicar que los conductores ajustan sus velocidades de manera más frecuente a medida que la variabilidad de la velocidad aumenta, lo cual induce en ellos más posibilidad de error en mantener la distancia de seguridad entre vehículos (Lee et al., 2002).
- La presencia de lluvia en tramos con curvas, es un factor que aumenta el riesgo de accidentes leves en DP pero graves en DT (“Sin Heridos y Bloqueo Mayor”). El pavimento húmedo ante maniobras de conducción realizadas por parte de los conductores como si fueran condiciones climáticas favorables, puede generar resbalamiento de los vehículos bloqueando más de una pista.
- En tramos con rampas de acceso (pista de entrada) o con gradiente positivo (pendiente en subida) disminuye la probabilidad de ocurrencia de los accidentes de tipo “Con Heridos y Bloqueo Mayor” e “Sin Heridos y Bloqueo Mayor”, respectivamente. Una pista de entrada puede generar precaución en

los conductores con los vehículos que están por entrar a la autopista. En cambio, una pendiente en subida puede generar velocidades menores, y por ende, menor riesgo.

- Algunas interacciones de las variables, muestran influencia en la ocurrencia de accidentes, y como era de esperar, en la mayoría de éstas actúan los Estados de Tráfico. Por ejemplo, la presencia de una pendiente de bajada combinada con estados de transición en promedio duplica las probabilidades de accidente.
- Otros factores que representan variables temporales indican que la ocurrencia de algunos accidentes es menos probable en días de semana mientras haya luz solar.

Se recuerda que este tipo de resultados están relacionados con factores influyentes previos a la ocurrencia de accidentes, y éstos son los que conciernen a los operadores, ya que son los que ellos pueden observar, y en algunos casos, controlar.

Dado lo anterior, a continuación se presentan algunas propuestas, recomendaciones y posibles direcciones para investigaciones futuras.

6.4 Recomendaciones y Trabajo Futuro

Con estos resultados, se pueden estudiar posibles soluciones, o bien, estrategias de asistencia y prevención que los operadores de autopistas podrían emplear para prevenir accidentes o asistirlos de mejor forma. Se propone, como simple sugerencia, la posibilidad de estudiar sistemas de prevención, tales como: sistemas de *ramp metering* y sistemas de límites de velocidad variable en función de los Estados de Tráfico. Estudios anteriores favorecen el uso de este sistema inteligente de transporte, ya que disminuye el riesgo de ocurrencia de accidentes.

También, con el uso de paneles de mensaje variable, se pueden mostrar mensajes para prevenir a los conductores de zonas o condiciones de operación de tráfico riesgosas. En fin, dar información al usuario para prevenir accidentes o inducir forma de conducción, que sean más seguras.

Para futuras investigaciones, se mencionan las siguientes recomendaciones:

- Estudiar el mismo fenómeno con datos obtenidos de otras autopistas urbanas presentes en Santiago de Chile, además de la Autopista Central. Por ejemplo, Costanera Norte, Vespucio Norte y Vespucio Sur.
- Considerar otro tipo de definición para los tipos de accidentes, para capturar otros efectos causados a partir de ellos. Es decir, otra manera que permita valorar las externalidades negativas que los accidentes viales producen. Por ejemplo, definir rangos de demora reportada para capturar el efecto que los accidentes producen al causar congestión medida por la demora.
- Estudiar factores influyentes en la ocurrencia de accidentes, en redes viales urbanas o interurbanas. No solo autopistas.
- Evaluar el efecto del posible uso de sistemas inteligentes de transporte que buscan prevenir la ocurrencia de accidentes, por medio de la simulación.

En cuanto a los resultados, se recomienda:

- Estudiar los factores de tráfico influyentes en la ocurrencia de accidentes, con el uso de datos obtenidos a partir de otros medios de recolección. Como por ejemplo, el uso de los pórticos de peaje.
- Considerar alternativas de medición de otro tipo de variables, además del tráfico, que provean mejor calidad de resolución de los datos. Como por ejemplo, mejor resolución de datos climáticos como intensidad de la lluvia, viento y neblina, o bien, información más detallada de la geometría de infraestructura vial, como valores de gradiente y radios de curvatura.
- Explorar el uso de otros modelos de estimación de probabilidades o predicción de ocurrencias de accidentes.
- Evaluar la posible inclusión de este tipo de resultados en la etapa de diseño de proyectos de infraestructura vial, con el fin de mitigar, a futuro, zonas riesgosas de accidentes.
- Considerar los resultados obtenidos, para estudiar formas de identificación de zonas riesgosas o *Hotspots*.

- Investigar sobre algoritmos o métodos eficientes para la actualización de los parámetros, con el uso de información nueva que las autopistas recolectan con baja frecuencia temporal por medio de las espiras magnéticas.

REFERENCIAS

Abdel-Aty, M. (2003). Analysis of driver injury severity levels at multiple locations using ordered probit models. *Journal of Safety Research*, 34(5), 597–603.

Abdel-Aty, M., & Abdelwahab, H. (2004). Modeling rear-end collisions including the role of driver's visibility and light truck vehicles using a nested logit structure. *Accident; analysis and prevention*, 36(3), 447–56.

Abdel-Aty, M., & Pande, A. (2004). Predicting Traffic Crashes Using Real-Time Traffic Speed Patterns. *Published in 10th World Conference on Transport Research, Istanbul, Turkey*.

Abdel-Aty, M., & Pande, A. (2005). Identifying crash propensity using specific traffic speed conditions. *Journal of Safety Research*, 36(1), 97–108.

Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, M. F., & Hsia, L. (2004). Predicting Freeway Crashes from Loop Detector Data by Matched Case-Control Logistic Regression. *Transportation Research Record: Journal of the Transportation Research Board* 1897.

Ahmed, M., Abdel-Aty, M., & Yu, R. (2012a). Assessment of Interaction of Crash Occurrence, Mountainous Freeway Geometry, Real-Time Weather, and Traffic Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2280(-1), 51–59.

Ahmed, M., Abdel-Aty, M., & Yu, R. (2012b). Bayesian Updating Approach for Real-Time Safety Evaluation with Automatic Vehicle Identification Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2280(-1), 60–67.

Ahmed, M., Huang, H., Abdel-Aty, M., & Guevara, B. (2011). Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway. *Accident; analysis and prevention*, 43(4), 1581–9.

Américo Vespucio Norte (2012). *Memoria Anual*

Américo Vespucio Sur (2012). *Memoria Anual*

Anastasopoulos, P. C., & Mannering, F. L. (2011). An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data. *Accident; analysis and prevention*, 43(3), 1140–7.

Autopista Central (2012). *Memoria Anual*

Bham, G. H., Javvadi, B. S., & Manepalli, U. R. R. (2012). Multinomial Logistic Regression Model for Single-Vehicle and Multivehicle Collisions on Urban U . S . Highways in Arkansas. *Journal of Transportation Engineering*, (June), 786–797.

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models. In *3rd Swiss Transportation Research Conference*. Ascona, Switzerland.

Christoforou, Z., Cohen, S., & Karlaftis, M. G. (2011). Identifying crash type propensity using real-time traffic data on freeways. *Journal of Safety Research*, 42(1), 43–50.

Christoforou, Z., Cohen, S., & Karlaftis, M. G. (2012). Integrating Real-Time Traffic Data in Road Safety Analysis. *Procedia - Social and Behavioral Sciences*, 48, 2454–2463.

Chung, K., Ragland, D. R., Madanat, S., & Oh, S. M. (2009). The Continuous Risk Profile Approach for the Identification of High Collision Concentration Locations on Congested Highways. *Transportation and Traffic Theory 2009*.

Costanera Norte (2012). *Memoria Anual*

Eluru, N., & Bhat, C. R. (2007). A joint econometric analysis of seat belt use and crash-related injury severity. *Accident; analysis and prevention*, 39(5), 1037–49.

Golob, T. F., & Recker, W. W. (2003). Relationships Among Urban Freeway Accidents, Traffic Flow, Weather and Lighting Conditions. *Journal of Transportation Engineering* 129, (August), 342–353.

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.

Golob, T. F., Recker, W. W., & Alvarez, V. M. (2004a). Freeway safety as a function of traffic flow. *Accident; analysis and prevention*, 36(6), 933–46.

Golob, T. F., Recker, W. W., & Alvarez, V. M. (2004b). Tool to Evaluate Safety Effects of Changes in Freeway Traffic Flow. *Journal of Transportation Engineering* 130, (April).

Greenshields, B. (1935). A study of traffic capacity. *Highway Research Board*, 448–477.

Haleem, K., & Abdel-Aty, M. (2010). Examining traffic crash injury severity at unsignalized intersections. *Journal of safety research*, 41(4), 347–57.

Hauer, E., & Persaud, B. (1997). Safety Analysis of Roadway Geometric and Ancillary Features. *Transportation Association of Canada*.

Hossain, M., & Muromachi, Y. (2012). A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. *Accident; analysis and prevention*, 45, 373–81.

Ivan, J. N. (2004). New Approach for Including Traffic Volumes in Crash Rate Analysis and Forecasting. *Transportation Research Record: Journal of the Transportation Research Board* 1897.

Jung, S., Qin, X., & Noyce, D. a. (2010). Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident; analysis and prevention*, 42(1), 213–24.

Khattak, A. J., Pawlovich, M. D., Souleyrette, R. R., & Hallmark, S. L. (2002). Factors Related to More Severe Older Driver Traffic Crash Injuries. *Journal of Transportation Engineering*, (June), 243–249.

Khorashadi, A., Niemeier, D., Shankar, V., & Mannering, F. (2005). Differences in rural and urban driver-injury severities in accidents involving large-trucks: an exploratory analysis. *Accident; analysis and prevention*, 37(5), 910–21.

Kockelman, K. M., & Kweon, Y.-J. (2002). Driver injury severity: an application of ordered probit models. *Accident; analysis and prevention*, 34(3), 313–21.

Kononov, J. (2002). Identifying Locations with Potential for Accident Reductions. *Transportation Research Record: Journal of the Transportation Research Board* 1784, 12(02), 153–158.

Kononov, J., Durso, C., Reeves, D., & Allery, B. K. (2012). Relationship Between Traffic Density, Speed, and Safety and Its Implications for Setting Variable Speed Limits on Freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 2280(-1), 1–9.

Kononov, J., Hersey, S., Reeves, D., & Allery, B. K. (2012). Relationship Between Freeway Flow Parameters and Safety and Its Implications for Hard Shoulder Running. *Transportation Research Record: Journal of the Transportation Research Board*, 2280(-1), 10–17.

Lave, C. A. (1985). Speeding, Coordination, and the 55 MPH Limit. *The American Economic Review*, 75(5), 1159–1164.

Lee, C., Hellinga, B., & Saccomanno, F. (2006). Evaluation of variable speed limits to improve traffic safety. *Transportation Research Part C: Emerging Technologies*, 14(3), 213–228.

Lee, C., Saccomanno, F., & Hellinga, B. (2002). Analysis of Crash Precursors on Instrumented Freeways. *Transportation Research Record: Journal of the Transportation Research Board* 1784, (02).

Lighthill, M. J., & Whitham, G. B. (1955). On Kinematic Waves. I. Flood Movement in Long Rivers. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 229(1178), 281–316.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, California: Sage Publications.

Malyshkina, N. V., & Mannering, F. (2008). Effect of Increases in Speed Limits on Severities of Injuries in Accidents. *Transportation Research Record: Journal of the Transportation Research Board*, 2083(-1), 122–127.

Martin, J.-L. (2002). Relationship between crash rate and hourly traffic flow on interurban motorways. *Accident; analysis and prevention*, 34(5), 619–29.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Economic Theory Mathematical Economics. Frontiers in Econometrics* (Zarembka., pp. 105–142). New York: Academic Press.

Oh, J.-S., Oh, C., Ritchie, S. G., & Chang, M. (2005). Real-Time Estimation of Accident Likelihood for Safety Enhancement. *Journal of Transportation Engineering*, 131(5), 358–363.

Ortúzar, J. de D. (2000). *Modelos Económicos de Elección Discreta*. Ediciones Universidad Católica de Chile.

Ortúzar, J. de D. y Willumsen, L.G. (2011). *Modelling Transport*. Cuarta Edición, John Wiley and Sons, Chichester.

Pande, A., & Abdel-Aty, M. (2009). A novel approach for analyzing severe crash patterns on multilane highways. *Accident; analysis and prevention*, 41(5), 985–94.

Park, J., & Oh, C. (2009). Relating Freeway Traffic Accidents to Inductive Loop Detector Data Using Logistic Regression. In *Paper Presented at the 4th IRTAD Conference, 16-17 September 2009, Seoul, Korea*. (pp. 223–231).

Quddus, M. a., Wang, C., & Ison, S. G. (2010). Road Traffic Congestion and Crash Severity: Econometric Analysis Using Ordered Response Models. *Journal of Transportation Engineering*, 136(5), 424–435.

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. a. (2011). The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. *Accident; analysis and prevention*, 43(5), 1666–76.

Schlesselman, J. (1982). *Case–Control Studies: Design, Conduct, Analysis*. Oxford University Press, New York.

TRB (2010). *Highway Capacity Manual, 2010*. Transportation Research Board of the National Academies, Washington, D.C.

Wang, C., Quddus, M. a, & Ison, S. G. (2009). Impact of traffic congestion on road accidents: a spatial analysis of the M25 motorway in England. *Accident; analysis and prevention*, 41(4), 798–808.

Wu, N. (2002). A new approach for modeling of Fundamental Diagrams. *Transportation Research Part A: Policy and Practice*, 36(10), 867–884.

Xu, C., Liu, P., Wang, W., & Li, Z. (2012). Evaluation of the impacts of traffic states on crash risks on freeways. *Accident; analysis and prevention*, 47, 162–71.

Yamamoto, T., & Shankar, V. N. (2004). Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. *Accident; analysis and prevention*, 36(5), 869–76.

Yang, Z., Zhibin, L., Pan, L., & Liteng, Z. (2011). Exploring contributing factors to crash injury severity at freeway diverge areas using ordered probit model. *Procedia Engineering*, 21, 178–185.

Ye, F., & Lord, D. (2013). Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. *Analytic Methods in Accident Research*, 1–14.

Yeo, H., Jang, K., Skabardonis, A., & Kang, S. (2013). Impact of traffic states on freeway crash involvement rates. *Accident; analysis and prevention*, 50, 713–23.

Zheng, Z. (2012). Empirical Analysis on Relationship between Traffic Conditions and Crash Occurrences. *Procedia - Social and Behavioral Sciences*, 43, 302–312.

Zheng, Z., Ahn, S., & Monsere, C. M. (2010). Impact of traffic oscillations on freeway crash occurrences. *Accident; analysis and prevention*, 42(2), 626–36.

Zúñiga, F. (2011). *Estimación y Predicción de Matrices Dinámicas de Viaje sobre un Corredor de Transporte Público, utilizando datos Históricos e Información en Tiempo Real*. Pontificia Universidad Católica de Chile.

ANEXOS

Anexo A: Otros Modelos de Predicción

A-1 Principal Component Analysis (PCA) y Nonlinear Canonical Correlation Analysis (NLCCA)

El método de las Componentes Principales (PCA) busca eliminar los problemas de multicolinealidad que existen entre variables que están muy correlacionadas entre sí. Define factores que son funciones lineales de las variables originales, tratando de retener la mayor cantidad de varianza.

Por otro lado, el método *Nonlinear Canonical Correlation Analysis* (NLCCA) es una forma de análisis de correlación canónica no lineal en donde variables categóricas son escaladas óptimamente como un componente integral para encontrar combinaciones lineales de variables con las correlaciones más altas entre ellas.

Comentarios

A pesar de que PCA soluciona el problema de multicolinealidad, los estados de tráfico están representados por una sola variable de tráfico al ser definidos discretizando una de las variables presentes en la relación fundamental, y por lo tanto, el problema de multicolinealidad deja de existir. Además, al utilizar PCA se pierde la posibilidad de interpretar los parámetros estimados de las variables originales, lo que constituye un problema no menor.

Como se menciona antes, NLCCA considera relaciones no lineales entre las variables. Sin embargo, es un modelo complejo en comparación al análisis de correlación canónica que solo considera relaciones lineales. No obstante, el uso de NLCCA muestra patrones claros al relacionar características de accidentes y condiciones de flujo dominante.

A-2 Redes Neuronales

Una explicación breve y simple de lo que son las redes neuronales se obtiene a partir de lo estipulado por Zúñiga (2011).

Las Redes Neuronales son modelos matemáticos de optimización no lineales del tipo “caja negra”, ya que son estudiadas desde el punto de vista de los datos de entrada y las salidas que produce.

La particularidad de estos modelos, y que hace mención a su nombre, es que están inspiradas en el comportamiento y capacidades del cerebro humano.

El modelo típico de Redes Neuronales asume que la salida que produce una neurona (nodo), corresponde a aplicar una función de activación $f(\cdot)$ sobre la suma ponderada de las entradas x_i por los pesos sinápticos w_i (pesos en los arcos). En la Figura 8-1: se presenta un esquema de una neurona artificial.

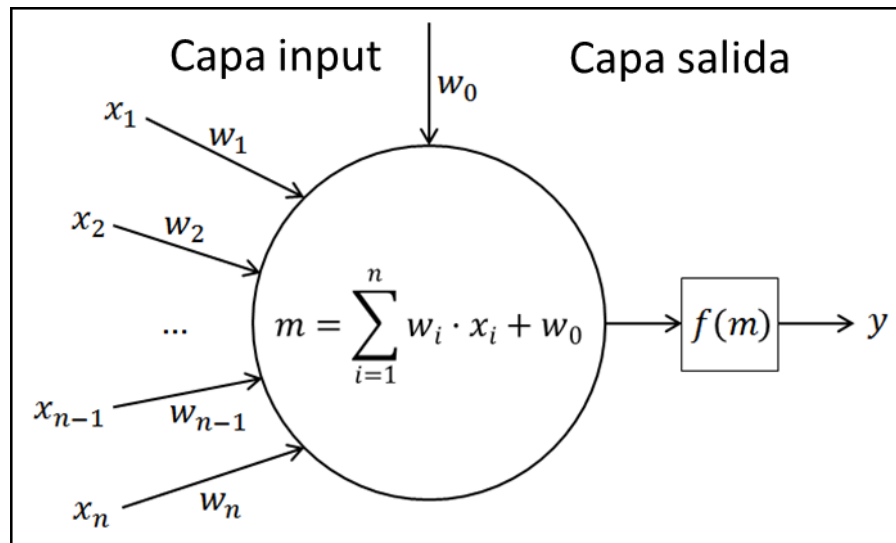


Figura 8-1: Esquema de una neurona artificial.

Comentarios

Las redes neuronales son un modelo bayesiano de aprendizaje adaptativo y aprenden a diferenciar patrones, por lo tanto no necesitan la especificación de un modelo *a priori* ni funciones de distribución de probabilidad. Sin embargo, es necesario definir previamente la forma funcional de la función de activación, lo cual en muchos casos no suele ser un trabajo trivial.

Además, la falta de reglas definitorias indica que hay muchos factores a tomar en cuenta: el algoritmo de aprendizaje, la arquitectura, el número de neuronas por capa, el número de capas, la representación de los datos, entre otros.

La distancia euclidiana entre neuronas, utilizada para definir los pesos sinápticos, no es estadísticamente satisfactoria, pues no toma en cuenta las diferencias a lo largo de los ejes, ni tampoco la posible correlación entre las variables de entrada. Algunos estudios consideran PCA para eliminar la correlación de las variables.

Al tratarse de una “caja negra”, no es posible ver, a diferencia de modelos estadísticos, las variables que fueron significativas y escogidas para estimar. Algunos consideran el uso de modelos de regresión lineal para identificar que variables son las más adecuadas como entrada a la red.

En el buen sentido, las redes permiten trabajar con datos de cualquier naturaleza, ya sean: lineales, no lineales, continuos, discretos, entre otros; y son muy tolerantes a información de entrada a la red con datos faltantes y mucho ruido.

Anexo B: Modelos Ordinales

Los modelos que consideran una escala ordinal para la variable dependiente, es decir, las categorías de ésta se ordenan naturalmente; se conocen como Modelos Ordinales.

La teoría del modelo Logit Ordinal (OLM) se describe en Long (1997). El OLM puede ser derivado de un modelo de medición en donde una variable latente y^* que va desde $-\infty$ hasta ∞ es asignada a una variable observada y . La variable y es considerada como el suministro de información incompleta sobre un y^* subyacente de acuerdo con la ecuación de medida:

$$y_i = m \quad \text{si } \tau_{m-1} \leq y_i^* < \tau_m \quad \text{para } m = 1 \text{ hasta } J \quad (8-1)$$

Los parámetros τ se denominan *thresholds* o puntos de corte. Las categorías extremas 1 y J son definidas por intervalos sin límites fijos $\tau_0 = -\infty$ y $\tau_J = \infty$. A modo de ejemplo y explicar la ecuación de medida, considere la variable dependiente que corresponde a niveles de siniestralidad de accidentes. Las categorías de respuesta son “Ilesos” (I), “Heridos” (H) y “Muerte” (M), es decir, $m = 1, 2, 3$. Asuma que esta variable ordinal está relacionada a una variable latente continua y^* que indica el grado de severidad de un accidente. La y observada está relacionada con y^* de acuerdo al modelo de medida siguiente:

$$y_i = \begin{cases} 1 \rightarrow I & \text{si } \tau_0 = -\infty \leq y_i^* < \tau_1 \\ 2 \rightarrow H & \text{si } \tau_1 \leq y_i^* < \tau_2 \\ 3 \rightarrow M & \text{si } \tau_2 \leq y_i^* < \tau_3 = \infty \end{cases} \quad (8-2)$$

La asignación de la variable latente a las categorías observadas se aprecia también en la Figura 8-2.

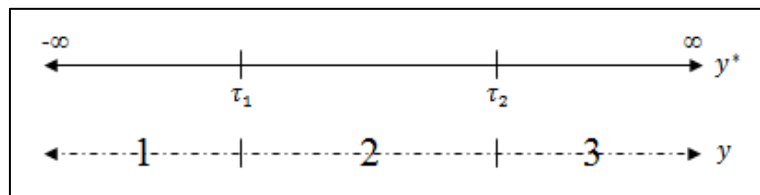


Figura 8-2: Esquema del Modelo de Medida.

El modelo estructural se expresa a través de la siguiente ecuación:

$$y_i^* = x_i\beta + \varepsilon_i \quad (8-3)$$

x_i es un vector fila con un 1 en la primera columna por el intercepto y la i -ésima observación para x_k en la columna $k + 1$. El vector de parámetros β es un vector columna para los parámetros estructurales con el primer elemento siendo el intercepto β_0 . ε_i corresponde al término de error con distribución de probabilidad Gumbel. Si la distribución fuese normal estaríamos hablando de un modelo Probit Ordinal (OPM).

El OLM, en términos de probabilidad para valores observados, se puede escribir de la siguiente manera:

$$P(y_i = m|x_i) = F(\tau_m - x_i\beta) - F(\tau_{m-1} - x_i\beta) \quad (8-4)$$

La Ecuación (8-4) es la probabilidad para cualquier respuesta m dado un x_i considerando que $F(\cdot)$ corresponde a la función de probabilidad acumulada de la distribución Gumbel. Si $y_i = 1$, el segundo término de la parte derecha de la Ecuación (8-4) se elimina, pues $F(\tau_0 - x_i\beta) = F(-\infty - x_i\beta) = 0$; y cuando se calcula la probabilidad de $y_i = J$ dado un x_i , el primer término del lado derecho de la Ecuación (8-4) es igual a 1, ya que $F(\tau_J - x_i\beta) = F(\infty - x_i\beta) = 1$.

Anexo C: Análisis de Correlación

Para realizar el Análisis de Correlación se utilizaron las estaciones “Río Mapocho en Rinconada Maipu” y “Cerro Calán” también ubicadas en la Región metropolitana en las coordenadas UTM Este: 331031, UTM Norte: 6291857, y UTM Este: 356880, UTM Norte: 6306490; respectivamente. Esto significa que la primera estación se encuentra a 16,85 kilómetros en distancia lineal de nuestra estación y la segunda a 11,65 kilómetros.

El análisis propiamente tal se hizo a partir de variables dicotómicas que indican si existe lluvia o no mediante la precipitación horaria registrada en cada estación el año 2012, lo que arrojó los siguientes resultados:

- La correlación entre las estaciones “Río Mapocho en Rinconada Maipu” y “Las Terrazas Oficinas DGA” fue de 65,29%.
- La correlación entre las estaciones “Cerro Calán” y “Las Terrazas Oficinas DGA” fue de 63,74%.

Considerando que las correlaciones no son menores, esto afirma nuestro supuesto anteriormente mencionado, ya que las distancias entre los extremos norte y sur de la autopista con respecto a la estación de Las Terrazas son muy cercanas a las distancias entre nuestra estación y las que se utilizaron para realizar este análisis.

Anexo D: Cálculo de V_f

Para obtener V_f se calcularon las velocidades para cada minuto en un punto de medida, para esto, se utilizó el promedio ponderado por el flujo de la velocidad, utilizando las velocidades medidas por los detectores pertenecientes a un punto de medida, en el mismo minuto.

Específicamente, la metodología para calcular V_f consiste en utilizar esta media ponderada de cada punto de medida para cada minuto y graficarlas en el tiempo. Se extrajeron las velocidades de ciertos puntos de medida, escogidos arbitrariamente del eje Ruta 5.

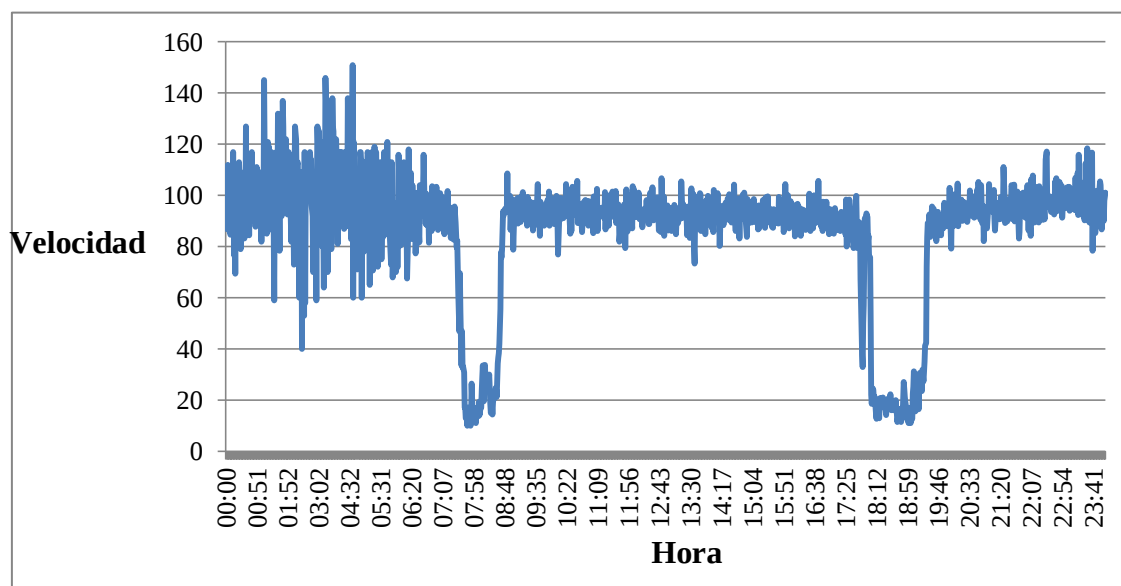


Figura 8-3: Gráfico de Velocidades medidas en PM007110-C1 el día 14/05/2012.

Los datos que se utilizaron corresponden a una semana de mediciones de 10 puntos de medida instalados en la autopista. La semana representativa escogida no considera feriados y va desde el día Lunes 14 de Mayo de 2012 hasta el Domingo 20 de Mayo 2012. En la Figura 8-3 se puede apreciar el perfil de velocidades en el tiempo, para un punto de medida específico en un día específico. La velocidad entre las 09:00 horas y las 17:30 horas se mantiene relativamente constante y con una variabilidad

pequeña, lo que constituye el criterio de extracción de observaciones de los otros puntos de medida y resto de los días.

Finalmente, con estas 38.854 observaciones, obtenidas a partir del criterio mencionado, se calculó el promedio de las velocidades para obtener 93 km/hr como la Velocidad de Flujo Libre y una desviación estándar de 6,59 km/hr. Además, se calculó la media de las velocidades para cada día y estas se aprecian en la Tabla 8-1.

Tabla 8-1: V_f por día de semana.

DIA	V_f
Lunes	91,85
Martes	91,67
Miércoles	91,63
Jueves	90,01
Viernes	89,97
Sábado	95,05
Domingo	97,45

Anexo E: Resultados de MNL 1:1 – 1:4

E-1 Modelo Binario

Tabla 8-2: Resultados de Estimación con 85% de muestra – Modelo Binario 1:1.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA	0,648	0,12
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA	1,180	0,02
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	0,072	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	0,051	0,00
ESTADO FINAL DE COLA	1,930	0,00
ESTADO CONGESTIÓN	1,010	0,00
ESTADO CUELLO DE BOTELLA	1,780	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3	0,0005	0,08
LLUVIA x CURVA	1,400	0,05
SEMANA	-0,287	0,06
CONSTANTE ACCIDENTE	-2,150	0,00
CONSTANTE NO-ACCIDENTE	0	--
Log-verosimilitud	-841,745	

Tabla 8-3: Resultados de Estimación con 85% de muestra – Modelo Binario 1:2.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA	0,740	0,02
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA	1,090	0,01
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	0,065	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	0,034	0,00
ESTADO FINAL DE COLA	1,890	0,00
ESTADO CONGESTIÓN	1,000	0,00
ESTADO CUELLO DE BOTELLA	1,860	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3	0,0006	0,02
SEMANA	-0,266	0,04
CONSTANTE ACCIDENTE	-2,650	0,00
CONSTANTE NO-ACCIDENTE	0	--
Log-verosimilitud	-1167,256	

Tabla 8-4: Resultados de Estimación con 85% de muestra – Modelo Binario 1:3.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA	0,659	0,02
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA	0,675	0,04
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	0,058	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	0,033	0,00
ESTADO FINAL DE COLA	2,010	0,00
ESTADO CONGESTIÓN	0,866	0,00
ESTADO CUELLO DE BOTELLA	1,940	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3	0,0005	0,02
SEMANA	-0,229	0,07
CONSTANTE ACCIDENTE	-2,940	0,00
CONSTANTE NO-ACCIDENTE	0	--
Log-verosimilitud	-1383,947	

Tabla 8-5: Resultados de Estimación con 85% de muestra – Modelo Binario 1:4.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA	0,574	0,02
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA	0,671	0,02
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS	0,056	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS	0,031	0,00
ESTADO FINAL DE COLA	2,020	0,00
ESTADO CONGESTIÓN	0,817	0,00
ESTADO CUELLO DE BOTELLA	1,900	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3	0,0005	0,04
SEMANA	-0,233	0,05
CONSTANTE ACCIDENTE	-3,120	0,00
CONSTANTE NO-ACCIDENTE	0	--
Log-verosimilitud	-1552,49	

E-2 Modelo 3 DP

Tabla 8-6: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:1.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos)	0,593	0,06
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,080	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,068	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos)	0,081	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,054	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos)	0,047	0,00
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos)	-0,260	0,06
ENTRADA (Con Heridos)	-0,333	0,02
ESTADO FINAL DE COLA (Sin Heridos)	2,310	0,00
ESTADO FINAL DE COLA (Con Heridos)	1,860	0,00
ESTADO CONGESTIÓN (Sin Heridos)	1,080	0,00
ESTADO CONGESTIÓN (Con Heridos)	0,707	0,03
ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,760	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos)	1,900	0,00
LLUVIA x CURVA (Sin Heridos)	1,360	0,07
LLUVIA x CURVA (Con Heridos)	1,520	0,05
CONSTATE (Con Heridos)	-2,670	0,00
CONSTATE (Sin Heridos)	-2,570	0,00
CONSTATE (No Accidente)	0	--
Log-verosimilitud	-1359,057	

Tabla 8-7: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:2.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos)	0,725	0,02
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,060	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,0598	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos)	0,074	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,037	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos)	0,030	0,00
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos)	-0,287	0,03
ENTRADA (Con Heridos)	-0,265	0,06
ESTADO FINAL DE COLA (Sin Heridos)	2,330	0,00
ESTADO FINAL DE COLA (Con Heridos)	1,810	0,00
ESTADO CONGESTIÓN (Sin Heridos)	1,140	0,00
ESTADO CONGESTIÓN (Con Heridos)	0,684	0,02
ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,880	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos)	1,960	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos)	0,0006	0,04
MERGE (Con Heridos)	0,700	0,09
CONSTATE (Con Heridos)	-3,120	0,00
CONSTATE (Sin Heridos)	-3,380	0,00
CONSTATE (No Accidente)	0	--
Log-verosimilitud	-1682,487	

Tabla 8-8: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:3.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos)	0,732	0,01
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos)	0,917	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,053	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos)	0,067	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,035	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos)	0,029	0,00
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos)	-0,293	0,02
ENTRADA (Con Heridos)	-0,248	0,07
ESTADO FINAL DE COLA (Sin Heridos)	2,410	0,00
ESTADO FINAL DE COLA (Con Heridos)	1,900	0,00
ESTADO CONGESTIÓN (Sin Heridos)	1,020	0,00
ESTADO CONGESTIÓN (Con Heridos)	0,544	0,06
ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,900	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos)	1,930	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos)	0,0006	0,04
MERGE (Con Heridos)	0,677	0,09
CONSTATE (Con Heridos)	-3,430	0,00
CONSTATE (Sin Heridos)	-3,660	0,00
CONSTATE (No Accidente)	0	--
Log-verosimilitud	-1896,760	

Tabla 8-9: Resultados de Estimación con 85% de muestra – Modelo 3 DP 1:4.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos)	0,719	0,01
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos)	0,951	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,050	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos)	0,066	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos)	0,034	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos)	0,027	0,00
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos)	-0,281	0,03
ENTRADA (Con Heridos)	-0,229	0,10
ESTADO FINAL DE COLA (Sin Heridos)	2,380	0,00
ESTADO FINAL DE COLA (Con Heridos)	1,880	0,00
ESTADO CONGESTIÓN (Sin Heridos)	0,979	0,00
ESTADO CONGESTIÓN (Con Heridos)	0,483	0,09
ESTADO CUELLO DE BOTELLA (Sin Heridos)	1,850	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos)	1,900	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos)	0,0005	0,06
MERGE (Con Heridos)	0,675	0,09
CONSTATE (Con Heridos)	-3,690	0,00
CONSTATE (Sin Heridos)	-3,860	0,00
CONSTATE (No Accidente)	0	--
Log-verosimilitud	-2064,503	

E-3 Modelo 3 DT

Tabla 8-10: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:1.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Bloqueo Menor)	-0,839	0,06
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	0,779	0,02
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,062	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,112	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,045	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,079	0,00
ESTADO FINAL DE COLA (Bloqueo Menor)	2,430	0,00
ESTADO FINAL DE COLA (Bloqueo Mayor)	1,610	0,00
ESTADO CONGESTIÓN (Bloqueo Menor)	1,700	0,00
ESTADO CONGESTIÓN (Bloqueo Mayor)	-0,703	0,16
ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	1,860	0,00
ESTADO CUELLO DE BOTELLA (Bloqueo Mayor)	1,880	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Bloqueo Mayor)	0,001	0,02
LLUVIA x CURVA (Bloqueo Mayor)	1,220	0,03
PENDIENTE EN SUBIDA (Bloqueo Mayor)	-0,291	0,10
SEMANA (Bloqueo Mayor)	-0,384	0,06
CONSTANTE (Bloqueo Menor)	-2,110	0,00
CONSTANTE (Bloqueo Mayor)	-4,470	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-1233,735	

Tabla 8-11: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:2.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Bloqueo Menor)	-0,891	0,02
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	0,92	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,056	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,107	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,029	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,065	0,00
ESTADO FINAL DE COLA (Bloqueo Menor)	2,420	0,00
ESTADO FINAL DE COLA (Bloqueo Mayor)	1,590	0,00
ESTADO CONGESTIÓN (Bloqueo Menor)	1,740	0,00
ESTADO CONGESTIÓN (Bloqueo Mayor)	-0,850	0,08
ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	1,850	0,00
ESTADO CUELLO DE BOTELLA (Bloqueo Mayor)	1,910	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Bloqueo Mayor)	0,001	0,01
LLUVIA x CURVA (Bloqueo Mayor)	0,966	0,07
PENDIENTE EN SUBIDA (Bloqueo Mayor)	-0,278	0,11
SEMANA (Bloqueo Mayor)	-0,386	0,05
CONSTANTE (Bloqueo Menor)	-2,57	0,00
CONSTANTE (Bloqueo Mayor)	-4,990	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-1556,607	

Tabla 8-12: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:3.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Bloqueo Menor)	-0,871	0,01
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	0,777	0,01
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,049	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,0998	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,029	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,064	0,00
ESTADO FINAL DE COLA (Bloqueo Menor)	2,500	0,00
ESTADO FINAL DE COLA (Bloqueo Mayor)	1,670	0,00
ESTADO CONGESTIÓN (Bloqueo Menor)	1,610	0,00
ESTADO CONGESTIÓN (Bloqueo Mayor)	-1,050	0,03
ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	1,880	0,00
ESTADO CUELLO DE BOTELLA (Bloqueo Mayor)	1,820	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Bloqueo Mayor)	0,001	0,01
LLUVIA x CURVA (Bloqueo Mayor)	0,828	0,11
PENDIENTE EN SUBIDA (Bloqueo Mayor)	-0,304	0,08
SEMANA (Bloqueo Mayor)	-0,362	0,06
CONSTANTE (Bloqueo Menor)	-2,890	0,00
CONSTANTE (Bloqueo Mayor)	-5,260	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-1770,839	

Tabla 8-13: Resultados de Estimación con 85% de muestra – Modelo 3 DT 1:4.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Bloqueo Menor)	-0,800	0,01
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	0,804	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,047	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,099	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Menor)	0,027	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Bloqueo Mayor)	0,063	0,00
ESTADO FINAL DE COLA (Bloqueo Menor)	2,470	0,00
ESTADO FINAL DE COLA (Bloqueo Mayor)	1,650	0,00
ESTADO CONGESTIÓN (Bloqueo Menor)	1,530	0,00
ESTADO CONGESTIÓN (Bloqueo Mayor)	-1,180	0,02
ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	1,850	0,00
ESTADO CUELLO DE BOTELLA (Bloqueo Mayor)	1,770	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Bloqueo Mayor)	0,001	0,02
LLUVIA x CURVA (Bloqueo Mayor)	0,812	0,11
PENDIENTE EN SUBIDA (Bloqueo Mayor)	-0,313	0,07
SEMANA (Bloqueo Mayor)	-0,361	0,06
CONSTANTE (Bloqueo Menor)	-3,130	0,00
CONSTANTE (Bloqueo Mayor)	-5,490	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-1937,980	

E-4 Modelo DP + DT

Tabla 8-14: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:1.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,587	0,09
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,628	0,13
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,060	0,00
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,160	0,05
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,062	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,101	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,061	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,118	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,045	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,110	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,038	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,058	0,00
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,395	0,01
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,491	0,04
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,480	0,00
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	1,070	0,02
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	2,020	0,00
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	1,900	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,750	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-2,200	0,05
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,906	0,01
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,800	0,00
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,650	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	2,100	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,530	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,002	0,00
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	1,660	0,02
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,775	0,00
SEMANA (Con Heridos y Bloqueo Mayor)	-0,416	0,10
CONSTANTE (Con Heridos y Bloqueo Menor)	-2,770	0,00
CONSTANTE (Con Heridos y Bloqueo Mayor)	-4,120	0,00
CONSTANTE (Sin Heridos y Bloqueo Menor)	-2,610	0,00
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-6,420	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-1722,540	

Tabla 8-15: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:2.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,698	0,04
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,759	0,05
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,140	0,00
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,160	0,04
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,056	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,094	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,055	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,111	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,029	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,0999	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,024	0,04
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,041	0,00
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,396	0,01
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,466	0,05
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,480	0,00
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	0,985	0,02
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	1,940	0,00
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	1,960	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,850	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-2,450	0,03
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,896	0,01
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,810	0,00
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,680	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	2,130	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,580	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,002	0,00
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	1,520	0,02
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,784	0,00
SEMANA (Con Heridos y Bloqueo Mayor)	-0,439	0,08
CONSTANTE (Con Heridos y Bloqueo Menor)	-3,230	0,00
CONSTANTE (Con Heridos y Bloqueo Mayor)	-4,570	0,00
CONSTANTE (Sin Heridos y Bloqueo Menor)	-3,060	0,00
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-6,940	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-2043,395	

Tabla 8-16: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:3.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,709	0,03
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,79	0,03
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,010	0,00
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	0,953	0,10
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,049	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,088	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,048	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,102	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,029	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,099	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,023	0,03
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,037	0,01
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,402	0,01
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,456	0,05
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,560	0,00
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	1,060	0,01
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	2,030	0,00
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	2,080	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,770	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-2,710	0,02
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,799	0,01
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,840	0,00
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,660	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	2,130	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,540	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,002	0,00
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	1,390	0,03
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,817	0,00
SEMANA (Con Heridos y Bloqueo Mayor)	-0,422	0,09
CONSTANTE (Con Heridos y Bloqueo Menor)	-3,550	0,00
CONSTANTE (Con Heridos y Bloqueo Mayor)	-4,830	0,00
CONSTANTE (Sin Heridos y Bloqueo Menor)	-3,370	0,00
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-7,180	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-2256,676	

Tabla 8-17: Resultados de Estimación con 85% de muestra – Modelo (DP+DT) 1:4.

Parámetro	Valor	Valor-p
PENDIENTE EN BAJADA x ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,708	0,03
PENDIENTE EN BAJADA x ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	-0,753	0,03
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,040	0,00
PENDIENTE EN BAJADA x ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	0,919	0,10
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,046	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,087	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,046	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 1 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,101	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Menor)	0,028	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Sin Heridos y Bloqueo Mayor)	0,099	0,00
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Menor)	0,022	0,04
DESVIACIÓN ESTÁNDAR VELOCIDAD ESTACIÓN 3 DE 5 MINUTOS PREVIOS (Con Heridos y Bloqueo Mayor)	0,034	0,01
DURANTE EL DÍA EN DÍA DE SEMANA (Con Heridos y Bloqueo Menor)	-0,392	0,01
ENTRADA (Con Heridos y Bloqueo Mayor)	-0,453	0,05
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	2,530	0,00
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	1,030	0,01
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	2,000	0,00
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	2,080	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	1,710	0,00
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	-2,900	0,01
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,767	0,02
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	1,800	0,00
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	1,630	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	2,110	0,00
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	1,520	0,00
DISTANCIA ENTRE ESTACIONES 1 Y 3 (Sin Heridos y Bloqueo Mayor)	0,002	0,00
LLUVIA x CURVA (Sin Heridos y Bloqueo Mayor)	1,440	0,03
PENDIENTE EN SUBIDA (Sin Heridos y Bloqueo Mayor)	-0,826	0,00
SEMANA (Con Heridos y Bloqueo Mayor)	-0,426	0,08
CONSTANTE (Con Heridos y Bloqueo Menor)	-3,800	0,00
CONSTANTE (Con Heridos y Bloqueo Mayor)	-5,060	0,00
CONSTANTE (Sin Heridos y Bloqueo Menor)	-3,620	0,00
CONSTANTE (Sin Heridos y Bloqueo Mayor)	-7,430	0,00
CONSTANTE (No Accidente)	0	--
Log-verosimilitud	-2423,266	

Anexo F: Errores Estándar Estimados

A modo de simplificación solo se presentan, a continuación, los errores estándar estimados de los parámetros asociados a las variables que representan los Estados de Tráfico, dado que en esta investigación se busca verificar la influencia de éstos en la ocurrencia de accidentes.

F-1 Modelo Binario

En la Tabla 8-18 se aprecia que a medida que la muestra crece, desde una razón Caso-Control de 1:1 hasta 1:5, todos los errores disminuyen.

Tabla 8-18: Errores Estándar Estimados con 85% de la muestra– Modelo Binario.

Parámetro	Errores Estándar Estimados				
	1:1	1:2	1:3	1:4	1:5
ESTADO FINAL DE COLA	0,328	0,242	0,221	0,204	0,196
ESTADO CONGESTIÓN	0,266	0,220	0,202	0,194	0,189
ESTADO CUELLO DE BOTELLA	0,298	0,242	0,222	0,207	0,198

F-2 Modelo 3 DP

En la Tabla 8-19 se aprecia que a medida que la muestra crece, desde una razón Caso-Control de 1:1 hasta 1:5, todos los errores disminuyen.

Tabla 8-19: Errores Estándar Estimados con 85% de la muestra – Modelo 3 DP.

Parámetro	Errores Estándar Estimados				
	1:1	1:2	1:3	1:4	1:5
ESTADO FINAL DE COLA (Sin Heridos)	0,235	0,189	0,174	0,163	0,159
ESTADO FINAL DE COLA (Con Heridos)	0,335	0,298	0,287	0,279	0,276
ESTADO CONGESTIÓN (Sin Heridos)	0,279	0,241	0,225	0,219	0,216
ESTADO CONGESTIÓN (Con Heridos)	0,332	0,299	0,289	0,286	0,286
ESTADO CUELLO DE BOTELLA (Sin Heridos)	0,303	0,266	0,249	0,241	0,237
ESTADO CUELLO DE BOTELLA (Con Heridos)	0,283	0,242	0,225	0,217	0,211

F-3 Modelo 3 DT

El modelo que considera como severidad solo la disrupción en el tráfico, muestra pequeñas diferencias en cuanto a la disminución de los errores estándar de los parámetros asociados a los Estados de Tráfico (ver Tabla 8-20).

El único parámetro que muestra un aumento en el error cuando se pasa desde 1:3 a 1:4 y vuelve a disminuir cuando la razón cambia desde 1:4 a 1:5, es aquel asociado al Estado “Congestión” de la alternativa que significa un accidente con “Bloqueo Mayor”. A pesar de esta oscilación, en lo general los errores se comportan de manera esperada.

Tabla 8-20: Errores Estándar Estimados con 85% de la muestra – Modelo 3 DT.

Parámetro	Errores Estándar Estimados				
	1:1	1:2	1:3	1:4	1:5
ESTADO FINAL DE COLA (Bloqueo Menor)	0,225	0,175	0,159	0,148	0,143
ESTADO FINAL DE COLA (Bloqueo Mayor)	0,314	0,283	0,273	0,266	0,265
ESTADO CONGESTIÓN (Bloqueo Menor)	0,387	0,316	0,278	0,259	0,256
ESTADO CONGESTIÓN (Bloqueo Mayor)	0,500	0,487	0,488	0,492	0,490
ESTADO CUELLO DE BOTELLA (Bloqueo Menor)	0,288	0,240	0,222	0,212	0,205
ESTADO CUELLO DE BOTELLA (Bloqueo Mayor)	0,334	0,298	0,283	0,277	0,275

F-4 Modelo DP + DT

En cuanto a los errores estándar de los parámetros asociados a las variables que representan los Estados de Tráfico, la gran mayoría presenta una disminución en los errores cuando la muestra aumenta de tamaño. Esto se puede apreciar en la Tabla 8-21.

Los casos en que la disminución de los errores estándar estimados no se da son para los parámetros de los Estados de Tráfico “Congestión” y “Cuello de Botella” de la alternativa “Sin Heridos y Bloqueo Mayor”.

En el primer caso, el error sufre un aumento de carácter despreciable cuando la muestra va desde 1:1 hasta 1:5, teniendo inicialmente un error de 1,110 y terminando en 1,150.

En segundo lugar, el error experimenta una disminución hasta llegar a la muestra de 1:4, pero vuelve a aumentar con la muestra de 1:5. Esto también se considera como un aumento despreciable, ya que ese aumento se debe a un aumento en el valor-p del

parámetro, pero no un aumento suficiente para que el parámetro deje de ser significativamente distinto de cero.

Tabla 8-21: Errores Estándar Estimados con 85% de la muestra – Modelo 5 Alternativas (DP+DT).

Parámetro	Errores Estándar Estimados				
	1:1	1:2	1:3	1:4	1:5
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Menor)	0,240	0,193	0,179	0,169	0,165
ESTADO FINAL DE COLA (Sin Heridos y Bloqueo Mayor)	0,449	0,433	0,428	0,422	0,426
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Menor)	0,364	0,331	0,322	0,315	0,313
ESTADO FINAL DE COLA (Con Heridos y Bloqueo Mayor)	0,342	0,311	0,300	0,293	0,291
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Menor)	0,358	0,319	0,293	0,280	0,278
ESTADO CONGESTIÓN (Sin Heridos y Bloqueo Mayor)	1,110	1,120	1,140	1,160	1,150
ESTADO CONGESTIÓN (Con Heridos y Bloqueo Menor)	0,348	0,328	0,321	0,319	0,319
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Menor)	0,314	0,274	0,258	0,252	0,246
ESTADO CUELLO DE BOTELLA (Sin Heridos y Bloqueo Mayor)	0,544	0,521	0,512	0,507	0,535
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Menor)	0,303	0,265	0,251	0,244	0,240
ESTADO CUELLO DE BOTELLA (Con Heridos y Bloqueo Mayor)	0,389	0,356	0,343	0,338	0,334