



Universidad de Concepción
Dirección de Postgrado
Facultad de Ingeniería -Programa de Magíster en Ciencias de la Ingeniería con mención
en Ingeniería Civil

Modelación de Factores Determinantes de Accidentes en Intersecciones Urbanas

Tesis para optar al grado de Magíster en Ciencias de la Ingeniería con
mención en Ingeniería Civil

RENÉ ALFREDO DUFEU FULLER
CONCEPCIÓN-CHILE
2019

Profesor Guía: Tomás Echaveguren Navarro
Dpto. de Ingeniería Civil, Facultad de Ingeniería
Universidad de Concepción

RESUMEN

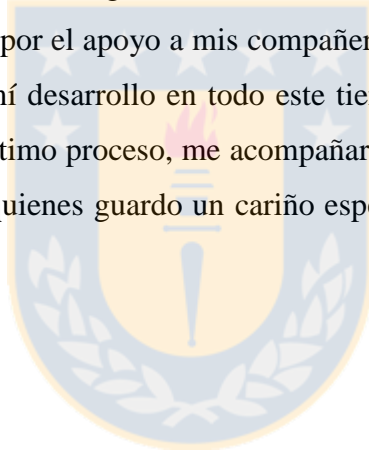
Los accidentes de tráfico son un problema creciente en la realidad de las ciudades. Estos eventos, cuyas causales son de diversa índole, generan impactos negativos en el sistema de transporte, en el sistema de actividades, en los involucrados y también en los espectadores, ocasionando daños materiales, físicos y psicológicos. Hoy en día más de la mitad de los accidentes urbanos ocurren en las intersecciones, pues son puntos complejos donde convergen distintos tipos de usuarios y modos de transporte. El objetivo principal es identificar los factores relacionados con las características de infraestructura, operación y entorno que explican la ocurrencia y el número de accidentes en las intersecciones urbanas.

Para lo anterior se realizó un acabado análisis del estado del arte, de tal modo de identificar y conceptualizar los factores que comúnmente son utilizados para estudiar la ocurrencia de accidentes en intersecciones urbanas. Con esta base teórica se procedió elaborar la metodología del análisis. Luego, se compiló y generó la base de datos relacionada con las intersecciones y los accidentes, la cual se obtuvo realizando mediciones en terreno, consolidando información de diversos conteos vehiculares y recopilando los datos de accidentes del Gran Concepción, a través de Carabineros. Finalmente, se obtuvieron las relaciones entre variables explicativas y la ocurrencia de accidentes, a través de un análisis de efectos sobre la media, un análisis bi-variado y un análisis utilizando modelos de regresiones lineales y no lineales.

De los factores que se identificaron como contribuyentes en la causalidad de accidentes en las intersecciones urbanas del Gran Concepción estudiadas, se obtuvo que aquellas variables relacionadas con la operación son las que tienen una mayor significancia estadística con la ocurrencia de accidentes. En segundo lugar, una cantidad menor de variables relacionadas con la infraestructura también resultaron significativas con respecto al número de accidentes. Finalmente, las variables atribuibles al entorno no resultaron ser significativas estadísticamente en los distintos análisis de relaciones con la ocurrencia de accidentes en intersecciones urbanas, en los modelos estimados y analizados no resultaron ser variables significativas.

AGRADECIMIENTOS

Se agradece a la Unidad Operativa de Control de Tráfico (UOCT), a Secretaría de Planificación de Transporte (SECTRA) y al Laboratorio GESITRAN Biobío, por sus aportes significativos en la recopilación de antecedentes y datos que permitieron la realización de este estudio. Al profesor guía de esta investigación, Tomás Echaveguren, quien de forma perseverante me apoyó durante el desarrollo de esta investigación. Al profesor Alejandro Tudela, quien aportó constructivamente al estudio. A mi familia, Emilio, Roxana, Mathías y Carolina, por ser un pilar fundamental, entregándome las herramientas y los valores necesarios para mi formación no sólo académica, sino que también personal, por lo cual les agradezco profundamente lo que soy hoy en día. Agradezco a Carla su inmenso amor y su apoyo, constantemente entregado durante este largo proceso, no exento de problemas pero de los cuales gracias a todos he sido capaz de superarlos y ser una mejor persona y profesional. Gracias por el apoyo a mis compañeros y amigos, especialmente a quienes de algún modo aportaron en mí desarrollo en todo este tiempo. También debo darle las gracias especiales a quienes, en este último proceso, me acompañaron y apoyaron en mi lugar de trabajo, a mis amigos de Multiruta, a quienes guardo un cariño especial agradezco su apoyo y compañía siempre.



ÍNDICE DE CONTENIDOS

CAPÍTULO 1	INTRODUCCIÓN	1
1.1	Motivación.....	1
1.2	Hipótesis de investigación.....	1
1.3	Objetivo general.....	2
1.4	Objetivos específicos.....	2
1.5	Metodología de trabajo.....	2
1.6	Principales resultados y conclusiones.....	3
1.7	Organización del informe.....	4
CAPÍTULO 2	ANÁLISIS DE MODELOS DE ACCIDENTES	5
2.1	Introducción.....	5
2.2	El concepto de accidente.....	5
2.3	Modelos de accidentes.....	5
2.4	Caracterización de los casos de estudio de accidentes analizados.....	10
2.5	Variables dependientes en los modelos.....	12
2.6	Variables influyentes en la ocurrencia de accidentes.....	13
2.7	Conclusiones.....	19
CAPÍTULO 3	DISEÑO DE UN EXPERIMENTO PARA ANALIZAR LOS ACCIDENTES	20
3.1	Introducción.....	20
3.2	Definición de variables.....	20
3.3	Matriz factorial.....	21
3.4	Conclusiones.....	23
CAPÍTULO 4	ELABORACIÓN DE LA BASE DE DATOS DE ACCIDENTES	24
4.1	Introducción.....	24
4.2	Elaboración de la base de datos.....	24

4.3	Definición de la variable dependiente.....	26
4.4	Definición de las variables independientes.....	26
4.5	Obtención de los flujos vehiculares	29
4.6	Caracterización de las intersecciones.....	30
4.7	Caracterización de accidentes	31
4.8	Conclusiones	36
CAPÍTULO 5 ANÁLISIS ESTADÍSTICO		38
5.1	Introducción	38
5.2	Procedimiento de análisis	38
5.3	Estadística descriptiva de las variables dependientes	39
5.4	Estadística descriptiva de las variables independientes.....	41
5.5	Análisis de valores extremos	46
5.6	Ajuste de distribuciones de la variable dependiente.....	47
5.7	Análisis de los efectos principales sobre los accidentes	48
5.8	Análisis bi-variado.....	50
5.9	Conclusiones	57
CAPÍTULO 6 MODELACION DE ACCIDENTES		59
6.1	Metodología de análisis	59
6.2	Modelación lineal	62
6.3	Modelación no lineal	71
6.4	Validación de modelos estimados.....	74
6.5	Comparación de modelos en la literatura.....	80
6.6	Conclusiones	83
CAPÍTULO 7 CONCLUSIONES		86
REFERENCIAS		91
ANEXOS		94

ANEXO 3.1: Matriz factorial conceptual del experimento	94
ANEXO 4.1: Periodización del Gran Concepción	95
ANEXO 4.2 Distribuciones espaciales de los accidentes de la muestra	97
ANEXO 5.1 Histogramas de frecuencias	99
ANEXO 5.2 Efectos sobre la media de accidentes de las variables analizadas	113
ANEXO 5.3 Resultados del análisis de efectos sobre la gran media de accidentes.....	138
ANEXO 6.1 Resultados de la modelación exploratoria, formas conceptuales y modelos en cada periodo de análisis	141
ANEXO 6.2 Resultados y salidas de modelaciones de Regresión Poisson	144



ÍNDICE DE TABLAS

Tabla 2.1 Niveles de severidad definidos según la escala KABCO.....	6
Tabla 2.2 Resultados de calibración del modelo	9
Tabla 2.3 Resumen de experimentos asociados a los modelos revisados	11
Tabla 2.4 Variables dependientes identificadas en modelos de accidentes en intersecciones urbanas	12
Tabla 2.5 Nomenclatura utilizada para definir las variables independientes	13
Tabla 2.6 Variables de infraestructura utilizadas en modelos de accidentes.....	15
Tabla 2.7 Variables de operación significativas.....	17
Tabla 2.8 Variables de entorno que explican la ocurrencia de accidentes en intersecciones.....	18
Tabla 3.1 Definición de variable dependiente del experimento	20
Tabla 3.2 Definición de variables independientes de infraestructura.....	20
Tabla 3.3 Definición de variables independientes relacionadas al entorno.....	21
Tabla 3.4 Definición de variables independientes relacionadas a la operación	21
Tabla 4.1 Bases de datos del estudio.	25
Tabla 4.2 .Variables del entorno de las intersecciones obtenidas para la base de datos	26
Tabla 4.3 .Variables de infraestructura de las intersecciones obtenidas para la base de datos	27
Tabla 4.4 .Variables de la operación de las intersecciones obtenidas para la base de datos	28
Tabla 4.5 Tasas de crecimiento vehicular adoptadas en este estudio.....	29
Tabla 4.6 Cantidad de intersecciones por comuna y tipo en base de datos	30
Tabla 4.7 Cantidad de accidentes por año de la población y de la muestra	32
Tabla 4.8 Accidentes por comunas en la BD del Gran Concepción y en la muestra	33
Tabla 4.9 Cantidad de accidentes por tipo de intersección en las seis bases de datos	36
Tabla 4.10 Promedio de accidentes por tipo de intersección en las seis bases de datos	36
Tabla 5.1 Estadística descriptiva de la variable dependiente en cada base de datos	41
Tabla 5.2 Frecuencias de las variables del tipo <i>dummy</i> de la base de datos.....	41
Tabla 5.3 Número de accidentes de intersecciones identificadas como valores extremos.....	47
Tabla 5.4 Composición inicial y post-análisis de valores extremos de las bases de datos.....	47
Tabla 5.5 Ajustes de distribuciones para la variable de accidentes.....	48
Tabla 5.6 Grupos de variables para el análisis de efectos principales sobre la media	49

Tabla 5.7 Correlación de Pearson del número de accidentes con las variables continuas en Día completo, Diurno y FP	51
Tabla 5.8 Correlación de Pearson del número de accidentes con las variables continuas en PM, PMD y PT.....	51
Tabla 5.9 Media de accidentes de cada nivel de cada variable dicotómica.....	53
Tabla 5.10 Significancia estadística según la prueba t-Student para cada variable de dos niveles.....	54
Tabla 5.11 Resultados del ANOVA de las variables de más de dos niveles.....	55
Tabla 5.12 Resumen de variables significativas con el n° de accidentes	57
Tabla 6.1 Modelos analizados.	63
Tabla 6.2 R ² de los modelos más adecuados en cada uno de los periodos de análisis.....	65
Tabla 6.3 Resultados de la forma conceptual más adecuada, Modelo 8.	66
Tabla 6.4 Resultados de la modelación Poisson, con la forma conceptual del modelo 8 como base	67
Tabla 6.5 Ajustes del modelo excluyendo la variable de flujo de camiones.....	68
Tabla 6.6 Ajustes del modelo excluyendo el flujo de camiones y taxibuses.....	68
Tabla 6.7 Formas conceptuales y ajustes de los modelos no lineales analizados.	72
Tabla 6.8 Indicadores de comparación en la estimación de los accidentes con el modelo no lineal versus los datos observados.....	80
Tabla A.3.1 Matriz factorial conceptual.....	94
Tabla A.4.1 Resultados de la periodización realizada, definiendo 2 horas en cada periodo.....	96
Tabla A.5.1 Ranking de importancia por grupo de variables según el análisis de efectos de todas las variables por cada grupo sobre la media de accidentes.....	138
Tabla A.5.2 Observación del comportamiento gráfico de cada variable peatonal en su efecto sobre la media de accidentes por periodo.....	139
Tabla A.5.3 Observación del comportamiento gráfico de cada variable relacionada a los giros en su efecto sobre la media de accidentes por periodo.....	139
Tabla A.5.4 Observación del comportamiento gráfico de cada variable de diseño geométrico en su efecto sobre la media de accidentes por periodo	140
Tabla A.5.5 Observación del comportamiento gráfico de cada variable de visibilidad en su efecto sobre la media de accidentes por periodo.....	140

Tabla A.5.6 Observación del comportamiento gráfico de cada variable en su efecto sobre la media de accidentes por periodo	140
Tabla A.6.1 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 1 en los distintos periodos de análisis.....	141
Tabla A.6.2 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 2 en los distintos periodos de análisis.....	141
Tabla A.6.3 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 3 en los distintos periodos de análisis.....	142
Tabla A.6.4 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 4 en los distintos periodos de análisis.....	142
Tabla A.6.5 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 5 en los distintos periodos de análisis.....	142
Tabla A.6.6 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 6 en los distintos periodos de análisis.....	143
Tabla A.6.7 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 7 en los distintos periodos de análisis.....	143
Tabla A.6.8 Modelo inicial con variable N° Cruces peatonales (Máxima Verosimilitud)	144
Tabla A.6.9 Análisis de Desviación	144
Tabla A.6.10 Pruebas de Razón de Verosimilitud.....	144
Tabla A.6.11 Análisis de Residuos.....	144
Tabla A.6.12 Modelo inicial con variable N° Cruces peatonales sin Flujo Camiones (Máxima Verosimilitud).....	146
Tabla A.6.13 Análisis de Desviación	146
Tabla A.6.14 Pruebas de Razón de Verosimilitud.....	146
Tabla A.6.15 Análisis de Residuos.....	146
Tabla A.6.16 Modelo inicial con variable N° Cruces peatonales sin Flujo Camiones y Taxibuses (Máxima Verosimilitud).....	147
Tabla A.6.17 Análisis de Desviación	148
Tabla A.6.18 Pruebas de Razón de Verosimilitud.....	148
Tabla A.6.19 Análisis de Residuos.....	148
Tabla A.6.20 Modelo inicial con variable N° Paraderos (Máxima Verosimilitud).....	149

Tabla A.6.21 Análisis de Desviación	149
Tabla A.6.22 Pruebas de Razón de Verosimilitud.....	149
Tabla A.6.23 Análisis de Residuos.....	150
Tabla A.6.24 Modelo inicial con variable Visibilidad (Máxima Verosimilitud).....	151
Tabla A.6.25 Análisis de Desviación	151
Tabla A.6.26 Pruebas de Razón de Verosimilitud.....	151
Tabla A.6.27 Análisis de Residuos.....	151
Tabla A.6.28 Modelo inicial con variable Visibilidad sin Flujo Camiones (Máxima Verosimilitud)	152
Tabla A.6.29 Análisis de Desviación	153
Tabla A.6.30 Pruebas de Razón de Verosimilitud.....	153
Tabla A.6.31 Análisis de Residuos.....	153



ÍNDICE DE FIGURAS

Figura 4.1 Estructura de la base de datos completa.....	24
Figura 4.2 Distancia de visibilidad.....	29
Figura 4.3 Evolución temporal comparativa de accidentes en intersecciones urbanas	32
Figura 4.4 Accidentes en intersecciones urbanas por comuna en el periodo de estudio.....	33
Figura 4.5 Accidentes en intersecciones urbanas según severidad	34
Figura 4.6 Accidentes de la base de datos de la muestra según su severidad	35
Figura 5.1 Esquema de la metodología del análisis de las variables del experimento.....	38
Figura 5.2 Histograma de frecuencia de los accidentes según periodo	40
Figura 5.3 Distribución de frecuencia del número de ramas de acceso en las intersecciones.....	42
Figura 5.4 Distribución de frecuencia del número de pistas en la rama principal	42
Figura 5.5 Distribución de frecuencia del número de pistas en la rama secundaria.....	42
Figura 5.6 Distribución de frecuencia del ángulo de la intersección.....	42
Figura 5.7 Distribución de frecuencia de la densidad peatonal de la intersección	43
Figura 5.8 Distribución de frecuencia del número de fases del semáforo de la intersección.....	43
Figura 5.9 Distribución de frecuencia del nivel de visión de la intersección	43
Figura 5.10 Distribución de frecuencia de la visibilidad de la intersección.....	44
Figura 5.11 Distribución de frecuencia de los cruces peatonales de la intersección.....	44
Figura 5.12 Distribución de frecuencia de los cruces peatonales con refugio de la intersección	44
Figura 5.13 Distribución de frecuencia de tipo de zona de las intersecciones	45
Figura 5.14 Distribución de frecuencia de número de centros de salud de las intersecciones	45
Figura 5.15 Distribución de frecuencia de número de paraderos de las intersecciones.	45
Figura 5.16 Distribución de frecuencia de número de centros de educación	45
Figura 5.17 Diagramas de caja de accidentes en cada periodo	46
Figura 6.1 Esquema de la metodología de la modelación	59
Figura 6.2 Número de accidentes modelados versus observados con el modelo de regresión lineal	74
Figura 6.3 Número de accidentes modelados versus observados con el modelo de Poisson.....	75
Figura 6.4 Número de accidentes modelados versus observados con el modelo no lineal	75
Figura 6.5 Número de Accidentes en intersecciones de Chiguayante.....	76
Figura 6.6 Número de Accidentes en intersecciones de San Pedro de la Paz	77

Figura 6.7 Número de Accidentes en intersecciones de Talcahuano	77
Figura 6.8 Número de Accidentes en intersecciones de Concepción con flujo mayor a 30.000 (veq/día).....	78
Figura 6.9 Número de Accidentes en intersecciones de Concepción con flujo menor a 30.000 (veq/día).....	79
Figura 6.10 Número de accidentes modelados y observados en el modelo de Cunto <i>et.al.</i> (2014)	81
Figura 6.11 Número de accidentes estimados con el modelo de Vieira Gomes (2013) versus el número de accidentes observados.....	81
Figura 6.12 Número de accidentes según el modelo de Vieira Gomes (2013) versus según el Modelo Propio	82
Figura A.4.1 Flujo total durante el día laboral en M. Rodríguez con Sta. Sofía, Chiguayante	95
Figura A.4.2 Flujo total durante el día laboral en P. Aguirre Cerda con Diag. Biobío, San Pedro.....	95
Figura A.4.3 Flujo total durante el día laboral en Av. Paicaví con Av. O'Higgins, Concepción	95
Figura A.4.4 Intersecciones y accidentes en la comuna de Concepción	97
Figura A.4.5 Intersecciones y accidentes en la comuna de Talcahuano.....	97
Figura A.4.6 Intersecciones y accidentes en la comuna de San Pedro de la Paz	98
Figura A.4.7 Intersecciones y accidentes en la comuna de Chiguayante	98
Figura A.5.1 Distribución de frecuencia de los flujos por tipo de movimiento del periodo PM ..	99
Figura A.5.2 Distribución de frecuencia de los flujos por tipo de movimiento del periodo PMD	100
Figura A.5.3 Distribución de frecuencia de los flujos por tipo de movimiento del periodo PT .	101
Figura A.5.4 Distribución de frecuencia de los flujos por tipo de movimiento del periodo FP..	102
Figura A.5.5 Distribución de frecuencia de los flujos por tipo de movimiento del periodo diurno	103
Figura A.5.6 Distribución de frecuencia de los flujos por tipo de movimiento del día completo	104
Figura A.5.7 Distribución de frecuencia de los flujos por tipo de vehículo del periodo PM.....	105
Figura A.5.8 Distribución de frecuencia de los flujos por tipo de vehículo del periodo PMD ...	106
Figura A.5.9 Distribución de frecuencia de los flujos por tipo de vehículo del periodo PT	107
Figura A.5.10 Distribución de frecuencia de los flujos por tipo de vehículo del periodo FP	108
Figura A.5.11 Distribución de frecuencia de los flujos por tipo de vehículo del periodo diurno	109

Figura A.5.12 Distribución de frecuencia de los flujos por tipo de vehículo del día completo ..	110
Figura A.5.13 Distribución de frecuencia de los cuocientes de flujos del periodo PM	111
Figura A.5.14 Distribución de frecuencia de los cuocientes de flujos del periodo PMD	111
Figura A.5.15 Distribución de frecuencia de los cuocientes de flujos del periodo PT	111
Figura A.5.16 Distribución de frecuencia de los cuocientes de flujos del periodo FP	112
Figura A.5.17 Distribución de frecuencia de los cuocientes de flujos del periodo diurno.....	112
Figura A.5.18 Distribución de frecuencia de los cuocientes de flujos del día completo.....	112
Figura A.5.19 Efectos del flujo por movimientos sobre la media de accidentes en PM.....	113
Figura A.5.20 Efectos del flujo por tipología sobre la media de accidentes en PM	114
Figura A.5.21 Efectos de los cuocientes del flujo sobre la media de accidentes en PM.....	114
Figura A.5.22 Efectos de las variables peatonales sobre la media de accidentes en PM.....	115
Figura A.5.23 Efectos de los giros sobre la media de accidentes en PM	115
Figura A.5.24 Efectos del diseño geométrico sobre la media de accidentes en PM	116
Figura A.5.25 Efectos de la visibilidad sobre la media de accidentes, base de datos PM	117
Figura A.5.26 Efectos de otras variables sobre la media de accidentes, base de datos PM.....	117
Figura A.5.27 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos PMD	118
Figura A.5.28 Efectos principales de variables de flujo por tipo de vehículos, base de datos PMD	118
Figura A.5.29 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos PMD.....	119
Figura A.5.30 Efectos principales de variables relacionadas al peatón, base de datos PMD.....	119
Figura A.5.31 Efectos principales de variables relacionadas a los giros, base de datos PMD....	120
Figura A.5.32 Efectos principales de variables relacionadas al diseño geométrico, base de datos PMD	120
Figura A.5.33 Efectos principales de variables relacionadas a la visibilidad, base de datos PMD	121
Figura A.5.34 Efectos principales de variables varias, base de datos PMD.....	121
Figura A.5.35 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos PT	122

Figura A.5.36 Efectos principales de variables de flujo por tipo de vehículos, base de datos PT	122
Figura A.5.37 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos PT	123
Figura A.5.38 Efectos principales de variables relacionadas al peatón, base de datos PT.....	123
Figura A.5.39 Efectos principales de variables relacionadas a los giros, base de datos PT.....	124
Figura A.5.40 Efectos principales de variables relacionadas al diseño geométrico, base de datos PT	124
Figura A.5.41 Efectos principales de variables relacionadas a la visibilidad, base de datos PT.	125
Figura A.5.42 Efectos principales de variables varias, base de datos PT.....	125
Figura A.5.43 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos FP	126
Figura A.5.44 Efectos principales de variables de flujo por tipo de vehículos, base de datos FP	126
Figura A.5.45 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos FP	127
Figura A.5.46 Efectos principales de variables relacionadas al peatón, base de datos FP.....	127
Figura A.5.47 Efectos principales de variables relacionadas a los giros, base de datos FP.....	128
Figura A.5.48 Efectos principales de variables relacionadas al diseño geométrico, base de datos FP	128
Figura A.5.49 Efectos principales de variables relacionadas a la visibilidad, base de datos FP.	129
Figura A.5.50 Efectos principales de variables varias, base de datos FP.....	129
Figura A.5.51 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos diurno	130
Figura A.5.52 Efectos principales de variables de flujo por tipo de vehículos, base de datos diurno	130
Figura A.5.53 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos diurno	131
Figura A.5.54 Efectos principales de variables relacionadas al peatón, base de datos diurno....	131
Figura A.5.55 Efectos principales de variables relacionadas a los giros, base de datos diurno..	132

Figura A.5.56 Efectos principales de variables relacionadas al diseño geométrico, base de datos diurno.....	132
Figura A.5.57 Efectos principales de variables relacionadas a la visibilidad, base de datos diurno	133
Figura A.5.58 Efectos principales de variables varias, base de datos diurno.....	133
Figura A.5.59 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos completa	134
Figura A.5.60 Efectos principales de variables de flujo por tipo de vehículos, base de datos completa	134
Figura A.5.61 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos completa.....	135
Figura A.5.62 Efectos principales de variables relacionadas al peatón, base de datos completa	135
Figura A.5.63 Efectos principales de variables relacionadas a los giros, base de datos completa	136
Figura A.5.64 Efectos principales de variables relacionadas al diseño geométrico, base de datos completa	136
Figura A.5.65 Efectos principales de variables relacionadas a la visibilidad, base de datos completa	137
Figura A.5.66 Efectos principales de variables varias, base de datos completa.....	137

CAPÍTULO 1 INTRODUCCIÓN

1.1 Motivación

En la actualidad los accidentes de tráfico son un problema creciente en la realidad del país, de la región y ciertamente, del Gran Concepción. De los accidentes que ocurren día a día, el 60% se desarrollan en zonas urbanas. Estos eventos, cuyas causas son de diversa índole, generan impactos negativos en el sistema de transporte, en el sistema de actividades y en los involucrados ocasionando daños materiales, físicos y psicológicos. Por este motivo es importante entender cómo se originan y así intentar evitarlos y/o mitigarlos. Ahora bien y según los registros, es en las intersecciones urbanas donde ocurre la mayor cantidad de accidentes de tránsito, debido a que son puntos complejos, donde convergen vehículos de diversos tipos además de peatones. Asimismo en las intersecciones existe interacción entre distintos modos de transporte, entre distintas clases de usuarios y existen muchos movimientos que se enfrentan. En este sentido, es común ver diseños y regulaciones que buscan ordenar dichas interacciones de modo tal de contribuir a la seguridad vial. No obstante, aún existen importantes desafíos para la academia y los agentes de planificación y de diseño urbano en cuanto al entendimiento y a la búsqueda de los factores que causan y promueven la ocurrencia de accidentes de tránsito. Es por esto que en este trabajo de investigación se busca dar una primera aproximación a la comprensión y a la determinación de los patrones de comportamiento y de las variables que explican la ocurrencia de accidentes, así como también el grado en que éstas puedan influir. De este modo sería posible pensar y generar diseños que maximicen la seguridad. Asimismo resultaría un importante aporte para la elaboración de mitigaciones certeras y *ad hoc* a cada contexto.

1.2 Hipótesis de investigación

Los accidentes en intersecciones urbanas son eventos que poseen componentes causales, es decir, pueden ser explicados por factores relacionados con la infraestructura, la operación y el entorno de las intersecciones.

1.3 Objetivo general

Identificar los factores que explican la ocurrencia y el número de accidentes en las intersecciones urbanas.

1.4 Objetivos específicos

- Identificar las variables relevantes y tipos de modelos para la estimación de accidentes en intersecciones urbanas.
- Definir un diseño experimental para poder analizar los accidentes en intersecciones urbanas del Gran Concepción.
- Elaborar una base de datos con la información necesaria para caracterizar los accidentes y las intersecciones.
- Realizar un análisis estadístico de los datos de accidentes en intersecciones urbanas.
- Estimar un modelo que estime el número potencial de accidentes en intersecciones urbanas en base a variables de entorno, infraestructura y operación.

1.5 Metodología de trabajo

La metodología de trabajo se dividió en cinco etapas. En principio, se realizó una revisión y clasificación de los modelos de accidentes en intersecciones urbanas. De esta etapa se obtuvo las principales variables explicativas de accidentes en intersecciones, los tamaños muestrales utilizados además de otros parámetros propios de la modelación de accidentes. En una segunda etapa se definieron las componentes del experimento para el análisis a realizar, es decir, las variables independientes y la variable explicada, la matriz factorial con todo el conjunto de variables independientes definiendo sus interacciones y niveles. Luego se identificaron los sitios de accidentabilidad (intersecciones) y se obtuvieron los datos de accidentes de la base de datos de Carabineros de Chile, la cual se reordenó y reestructuró. Como cuarta etapa, se caracterizaron las intersecciones urbanas del Gran Concepción. Esta caracterización se realizó en base a mediciones en terreno, catastros y mediciones vehiculares. Con la base de datos depurada, se analizó

estadísticamente las variables del experimento, se ajustaron distribuciones a las variables de accidentes y se realizó un análisis bi-variado, de tal manera de identificar patrones con los cuales poder obtener como resultado relaciones causales. Finalmente, en base a los resultados previos, se modeló el número de accidentes en intersecciones urbanas. Para ello se procedió a buscar las combinaciones de variables explicativas físicamente lógicas y adecuadas para obtener un modelo que maximice el ajuste y la simplicidad con el objetivo de estimar el número de accidentes por intersección. Dichos modelos fueron regresiones lineales, regresiones de Poisson y modelos no lineales.

1.6 Principales resultados y conclusiones

Se identificaron 13 variables relacionadas con la infraestructura, 19 variables de operación de la intersección y cuatro variables relacionadas con el entorno. Se encontró que los modelos utilizados en estudios previos se clasifican en modelos del tipo Probit, de Regresión y lineales generalizados. Además, se revisaron modelos con distribuciones cero-infladas cuyo objetivo es mejorar el ajuste por el exceso de ceros, considerando así la no ocurrencia de accidentes.

Se sintetizó una base de datos integrando la información de accidentes del Gran Concepción entre los años 2009 y 2013 (considerando los ocurridos en intersecciones urbanas) con la base de datos de las características de las intersecciones, la cual fue elaborada con mediciones en terreno y estudios de base que aportaron con la información requerida.

Al analizar individualmente la relación entre el número de accidentes y las variables, aquellas que mejor se relacionaron con la ocurrencia de accidentes fueron las de operación, de las cuales 16 resultaron significativas, entre ellas el flujo vehicular, la visibilidad y el número de fases de los semáforos. En segundo lugar, las variables relacionadas con la infraestructura, tales como el número de pistas, la presencia de mediana, el ángulo de la intersección, la presencia de pistas exclusivas de giro y el número de calles unidireccionales.

Finalmente, aquellas variables relacionadas con el entorno de la intersección no generan efectos significativos en el número de accidentes, siendo el número de paraderos la única variable

significativa. Con estos resultados se obtuvieron diversos modelos, tanto de regresiones lineales como no lineales. En este sentido, un modelo no lineal de forma exponencial es el que mejores ajustes presenta para la estimación de los accidentes, resultando como la variable con mayor peso estadístico e importancia, el flujo vehicular y en segundo plano, características físicas y operativas de las intersecciones, como lo fueron, el número de pistas, el ángulo de la intersección, la densidad peatonal, visibilidad, número de paraderos y número de cruces peatonales.

Otro resultado fue el nivel de ajustes de los modelos predictivos en base a las características de las intersecciones, los cuales alcanzaron una capacidad explicativa de 50% (R^2 ajustado). Esto quiere decir que el número de accidentes es posible relacionarlo con causas de infraestructura, operación y/o entorno de las intersecciones. No obstante existe una importante porción del fenómeno que no resulta del todo explicada por estas características y, por consiguiente, son otros aspectos y causalidades que también influyen en el número de accidentes, las cuales se encuentran fuera del alcance de esta investigación. Las cuales podrían ser características inherentes de los conductores, de los vehículos, así como condiciones aleatorias del entorno.

1.7 Organización del informe

Esta tesis consta de siete capítulos. En el capítulo dos se identifican los conceptos que gobiernan la modelación de accidentes en intersecciones, además de los distintos tipos de variables utilizadas, tamaños muestrales y los resultados en los distintos tipos de análisis. En el capítulo tres se propone el diseño del experimento, en el cual se definen las variables, dependiente e independientes, así como también el diseño factorial. Mientras que en el cuarto capítulo se presenta la obtención de las variables, la elaboración y el proceso de sintetizado de la base de datos. Posterior a esto, en el capítulo cinco, se realiza un análisis estadístico de los accidentes, de las variables dependientes e independientes y de sus relaciones, de tal modo de identificar los patrones de comportamiento. En el capítulo seis se presentan los resultados de la modelación de accidentes así como de la validación de los modelos escogidos como adecuados. Finalmente, en el capítulo siete se presentan las conclusiones de esta investigación.

CAPÍTULO 2 ANÁLISIS DE MODELOS DE ACCIDENTES

2.1 Introducción

En este capítulo se explican los conceptos fundamentales en los que se basa la investigación realizada. Para ello se presenta la definición del concepto de accidente, los tipos de modelos de accidentes en intersecciones urbanas encontrados y se identifican las variables que explican la ocurrencia de los accidentes.

2.2 El concepto de accidente

En Chile, la CONASET define los accidentes de tránsito como todo siniestro que ocurre en la vía pública y que tiene como consecuencia una o más personas muertas, lesionadas o ilesas y al menos, un vehículo en movimiento involucrado. Chías (2004) define a los accidentes de tránsito como eventos estocásticos, es decir, tienen una parte aleatoria y una parte determinística afirmando que se puede incidir sobre esta última para prevenir su ocurrencia. Asimismo, Gold (1998) había definido un accidente de tránsito de la misma manera, aseverando que en su origen existe una sucesión de evento cuyas componentes son tanto causales como aleatorias.

2.3 Modelos de accidentes

Los modelos para la estimación de accidentes en intersecciones urbanas se clasifican en tres grupos: modelos tipo Probit, de regresión y modelos lineales generalizados. A continuación se discute sobre cada uno de ellos.

2.3.1 Modelos tipo Probit

Estos modelos se utilizan para calcular la probabilidad de ocurrencia de un accidente y la probabilidad de ocurrencia asociada a un cierto nivel de severidad. Abdel-Aty y Keller (2004) utilizan el modelo de la Ecuación 2.1 para calcular la probabilidad de ocurrencia de un accidente con una cierta severidad.

$$\sum_{j=1}^j P_n(j) = F(\alpha_j - \beta_j X_n, \theta), \tag{2.1}$$

donde, F es la función de la distribución normal de probabilidad, θ es el parámetro de dispersión, X_n es un vector de características medibles, α_j la constante específica de cada categoría y β_j el vector de coeficientes estimables. $P_n(j)$ es la probabilidad de que el individuo n pertenezca a la categoría j de la variable dependiente, en otras palabras j representa las severidades de accidentes de tránsito. Cada nivel está definido según la escala de severidad KABCO de la Tabla 2.1.

Tabla 2.1 Niveles de severidad definidos según la escala KABCO (AASHTO, 2010)

Escala	Severidad	N° asociado en modelo
K	Accidentes fatal	0
A	Lesiones graves	1
B	Lesiones no-incapacitantes	2
C	Lesiones leves	3
O	Sin lesionados, solo daño material	4

Este tipo de modelación se ve afectada cuando muchas intersecciones presentan cero accidentes en los distintos periodos de análisis. Kumara y Chin (2003) utilizan modelos del tipo cero-inflado para mejorar la sobre dispersión inducida por el exceso de ceros. Así mismo existen modelos de mayor complejidad matemática, en donde a los modelos probit generalizados se incluye la dependencia espacial y la temporal, lo que permite analizar las correlaciones que existen debido a estas dos condiciones. Para esto, Castro *et al.* (2012) encontraron configuraciones espaciales entre intersecciones que generaban una tendencia común a obtener ciertos niveles de accidentes. También resultó significativo un término que agregaron a su modelo relacionado con la temporalidad, demostrándose la correlación debido a la repetición en el tiempo de los datos de accidentes de intersecciones.

Los accidentes se ajustan a distribuciones Poisson o Binomial Negativa (BN). La distribución BN tiene un mejor comportamiento que la Poisson, pues al existir mucha dispersión en los datos no se cumple el supuesto de la varianza igual a la media de una distribución de Poisson. Kumara y Chin (2003), Abdel-Aty y Keller (2004) y Greibe (2003) obtuvieron resultados que demuestran un mejor ajuste al utilizar la distribución BN. En lo que respecta a la sobre dispersión en este tipo de modelos, Kumara y Chin (2003) obtuvieron como resultado que el parámetro de dispersión resultaba siempre significativo, por lo cual a través de la distribución cero inflada mejoraron los ajustes en comparación a los modelos tradicionales.

2.3.2 Modelos de regresión

En los modelos de regresión las formas funcionales mayormente empleadas son multiplicativas, como por ejemplo en los modelos de Salifu (2003) o el de Cunto *et al.* (2014) (ver Ecuación 2.2).

$$E(\mu_i) = kQ^\alpha \exp(\sum \beta_j X_{ij}), \quad (2.2)$$

donde E es el número observado de accidentes, Q es el tráfico, k y α son coeficientes a estimar, β_j es el vector de coeficientes de la regresión asociados a cada variable explicativa j. X_{ij} son las variables explicativas asociadas a la geometría j de cada intersección i.

Por otro lado, el modelo de AASHTO (2010) consiste en un modelo de regresión que utiliza el método bayesiano. Se compone de SPFs (*Safety Performance Functions*), las que tienen una forma funcional exponencial y sus variables de entrada están establecidas en AASHTO (2010); la principal variable es el tráfico diario medio anual. Otra parte de su estructura radica en los factores modificadores de accidentes (CMF), los que se definen según AASHTO (2010), por tipo y por valor, para cada intersección. El modelo tiene un coeficiente de calibración C_x , el cual tiene por función adecuar las funciones SPFs al contexto en el cual se enmarque la modelación. El modelo para intersecciones queda representado de la siguiente manera (AASHTO, 2010).

$$N_{\text{predicted int}} = C_x (N_{\text{bi}} + N_{\text{pedi}} + N_{\text{bikei}}), \quad (2.3)$$

$$N_{\text{bi}} = N_{\text{spf}} (\text{CMF}_{1i} \times \text{CMF}_{2i} \times \dots), \quad (2.4)$$

$$\text{con, } N_{\text{spf}} = \exp(a + b \ln \text{AADT}_{\text{major}} + c \ln \text{AADT}_{\text{minor}}), \quad (2.5)$$

en donde, N_{spf} es el número de accidente predicho en condiciones base (para accidentes del tipo sólo-vehículo y vehículo-múltiple), AADT es el flujo medio diario anual en la calle principal y secundaria. N_{bi} es el número de accidentes estimados en una intersección excluyendo accidentes del tipo vehículo-peatón y vehículo-biciclo. Para la predicción de las colisiones con peatones y bicicletas involucrados existe una manera distinta de estimarlos. N_{pedi} es el número de accidente tipo vehículo-peatón en una intersección y N_{bikei} es lo mismo, pero para accidente entre vehículo y bicicletas.

Cunto *et al.* (2014) calibraron las ecuaciones de AASHTO (2010) en Brasil con información de flujos medios diarios y de accidentes de los años 2007 a 2009. Cuando existe información de accidentes, AASHTO (2010) recomienda utilizar el método de inferencia bayesiana (Fridstrom *et al.* 1995) para combinar los accidentes predichos con los observados, calculando el número de accidentes esperados según la Ecuación 2.6.

$$N_{\text{expected},i} = w_i N_{\text{predicted},i} + (1 - w_i) N_{\text{observed},i}, \quad (2.6)$$

$$w_i = \frac{1}{1+k (\sum \text{años} N_{\text{predicted},i})}, \quad (2.7)$$

en donde $N_{\text{expected},i}$ representa el número estimado de accidentes esperados, $N_{\text{observed},i}$ representa los accidentes observados en cada sitio i en el periodo de estudio y w_i es el ponderador de ajuste del modelo predictivo para cada sitio i . Finalmente se tiene que k es un parámetro de sobre dispersión y está asociado a las funciones SPF.

En el modelo de la AASHTO (2010), el aspecto derivado de la sobre dispersión de los datos es de vital importancia. Pues por un lado la capacidad de predicción del modelo está directamente relacionado con la sobre dispersión de la variable dependiente (número de accidentes), y además con las SPFs, las cuales están calibradas para Estados Unidos. Luego teóricamente, es ideal tener

las SPFs calibradas para cada contexto y lugar de utilización del modelo, así como también se debería calibrar el valor del parámetro k de sobre dispersión.

No obstante, Cunto *et al.* (2014) utilizaron tanto el valor k original de AASHTO (2010) para cada SPF como uno propio, es decir, recalibrando el parámetro k para cada modelo que estimaron, utilizando el procedimiento descrito por Sawalha y Sayed (2006). La Tabla 2.2 muestra los resultados obtenidos por Cunto *et al.* (2014) en Fortaleza, Brasil.

Tabla 2.2 Resultados de calibración del modelo (Cunto *et al.*, 2014)

	Semaforizadas	No Semaforizadas
Nº Intersecciones	45	46
Nº de accidentes observados	860	586
Nº de accidentes estimados	879	272
Coficiente de calibración C_x	0.98	2.15
Desviación media absoluta	2.2	2.4
Porcentaje de error medio absoluto	0.17	0.43

Para intersecciones semaforizadas, el coeficiente de calibración C_x fue de 0,98 y en intersecciones no semaforizadas de 2,15. Esto significa que el modelo de AASHTO (2010) predice de mejor manera los accidentes en intersecciones semaforizadas.

Cunto *et al.* (2014) observaron una sobre estimación de accidentes en sectores donde existían pocos y, al contrario, en sectores con un número de accidentes elevado, su modelo sub-estimó el número de accidentes.

Tasic y Porter (2016) buscaron definir la relación entre infraestructura multimodal de transporte y nivel de accidentes en zonas específicas de una ciudad. Para captar correlaciones espaciales a priori, y así evitar problemas en la modelación utilizaron el índice de Moran (Anselin y Arribas-Bel, 2013). Los autores elaboraron modelos que estimaban el número de accidentes vehiculares (todas las severidades), no motorizados (todas las severidades) y los accidentes fatales vehiculares. Para esto utilizaron variables espaciales, geométricas y de operación.

Los resultados de ajuste demostraron que resultaba mejor utilizar los modelos de efectos fijos BN (FENB) o de efectos aleatorios BN (RENB) en desmedro de uno BN (Pseudo $R^2=0,087$ versus

0,081). Así mismo identificaron diferencias en la significancia de las variables presencia de líneas de buses y de ciclo-vías, entre otras. En sus modelos, las variables socio económicas no resultaron estadísticamente significativas. Del mismo modo, los indicadores relativos al uso de suelo tampoco fueron significativos.

2.3.3 Modelos lineales generalizados

Los modelos lineales generalizados (GLM) sirven para estimar correlaciones entre factores independientes con respecto a la variable que se busca explicar. El resultado que se obtiene es una ponderación de cada factor con respecto a la variable explicada.

Este tipo de análisis también requiere de una cantidad elevada de datos pues, como observó Vieira Gomes (2013), los resultados de significancias de algunos coeficientes asociados a variables cambiaron con respecto al análisis de Vieira Gomes *et al.* (2012). Esto debido a que en ambos trabajos se consideraron las mismas intersecciones, pero en la segunda investigación se aumentó la cantidad de accidentes considerados. Por ejemplo variables como el flujo peatonal, o la presencia de mediana en la rama principal, resultaron en primera instancia no significativas en Vieira Gomes (2012), pero sí en Vieira Gomes (2013).

Mitra y Washington (2007) concluyeron también sobre la importancia de disponer de un elevado número de intersecciones con diferentes configuraciones para así poder disponer de datos en todas las combinaciones de factores del GLM y luego obtener resultados más precisos.

2.4 Caracterización de los casos de estudios de accidentes analizados

En la Tabla 2.8 se presenta el resumen del análisis de los modelos identificados, en donde se describe el lugar del estudio, el número de intersecciones consideradas, el número de accidentes total en las intersecciones, el periodo de años analizado y se describe lo que fue modelado.

El número de intersecciones utilizado por los diferentes autores varía entre 44 (Vieira Gomes, 2012) y 832 (Abdel-Aty y Keller, 2004). El rango del número de accidentes de las bases de datos de los modelos varía entre 150 (Tasic y Porter, 2016) y 21,204 (Abdel-Aty y Keller, 2004).

Tabla 2.3 Resumen de experimentos asociados a los modelos revisados

Tipo de Modelo	Autor	Origen	N° de inter.	N° de Accidentes	Periodo	Años	Modelado
Probit	Abdel-Aty y Keller (2004)	Florida, EE.UU.	832	21204	2000 - 2001	2	Severidad de accidentes
	Greibe (2003)	Dinamarca	547	270	1987 a 1991	5	Tasa de accidentes por año para inter. T semaforizadas
	Greibe (2003)	Dinamarca	85	584	1987 a 1991	5	Tasa de accidentes por año para inter. T no semaforizadas
	Greibe (2003)	Dinamarca	154	1345	1987 a 1991	5	Tasa de accidentes por año para inter. X semaforizadas
	Greibe (2003)	Dinamarca	250	335	1987 a 1991	5	Tasa de accidentes por año para inter. X no semaforizadas
	Kumara y Chin (2003)	Singapur	104	809	1992 a 2000	9	Tasa de accidentes por año por inter. T semaforizadas
GLM	Chen y Xie (2016)	Massachusetts	48	688	2009 a 2012	4	N° de accidentes por año por inter. T semaforizadas
	Chen y Xie (2016)	Massachusetts	52	899	2009 a 2012	4	N° de accidentes por año por inter. X semaforizadas
	Vieira Gomes et. al. (2012)	Lisboa	44	147	2004 a 2007	4	N° de accidentes motorizados por año por interseccion T
	Vieira Gomes et. al. (2012)	Lisboa	50	343	2004 a 2007	4	N° de accidentes motorizados por año por interseccion X
	Vieira Gomes (2013)	Lisboa	44	182	2004 a 2007	4	N° de accidente por año por inter. T considerando accidentes peatonales
	Vieira Gomes (2013)	Lisboa	50	391	2004 a 2007	4	N° de accidente por año por intersección X considerando accidentes peatonales
	Mitra y Washington (2007)	Georgia	165	837	1996 - 1997	2	Frecuencia de accidentes por año por intersecciones
Regresión	Tasic y Porter (2016)	Chicago	-	9895	2005 a 2012	8	N° de accidente vehicular por sector por año
	Tasic y Porter (2016)	Chicago	-	1122	2005 a 2012	8	N° de accidente no motorizado por sector por año
	Tasic y Porter (2016)	Chicago	-	150	2005 a 2012	8	N° de accidente fatal por sector por año
	Cunto et. al. (2014)	Fortaleza	89	1657	2007 a 2009	3	N° de accidentes por inter. semaforizada por año
	Cunto et. al. (2014)	Fortaleza	92	1099	2007 a 2009	3	N° de accidentes por inter. no semaforizada por año
	Salifu (2003)	Ghana	57	354	1996 a 1998	3	Media de accidentes por intersecciones tipo T
	Salifu (2003)	Ghana	34	238	1996 a 1998	3	Media de accidentes por intersecciones tipo X

El periodo de años de accidentes considerados varía entre los dos y nueve años. No se recomienda utilizar más de cinco años de accidentes, puesto que en ese lapso de tiempo pueden existir cambios importantes en las características de las intersecciones. Además, si no se considera explícitamente la correlación temporal en el modelo, el error asociado a este sesgo aumenta a medida que se utilizan periodos con más años (Castro *et al.*, 2012).

En lo que respecta la modelación, es común que se segmenten los datos y generen modelos para cada clase de intersección (Chen y Xie, 2016), así como modelos para cada tipo de accidente según el tipo y/o severidad (Wang *et al.*, 2011).

2.5 Variables dependientes en los modelos

Se observaron dos grandes opciones de modelación. La primera y mayormente utilizada en la literatura es la estimación del número de accidentes en un periodo de tiempo fijo y por intersección. La segunda se refiere a la estimación por severidad de accidentes, en donde se estiman ya sea una proporción de accidentes por tipo de severidad por intersección o el número en cada intersección y por severidad.

Tabla 2.4 Variables dependientes identificadas en modelos de accidentes en intersecciones urbanas

Variabes Dependientes	Nº de accidentes por severidad por año y por intersección	Nº de accidentes por año y por intersección
Greibe (2003)	-	Continua
Kumara y Chin (2003)	-	Discreta (0,1,...,6)
Salifu (2003)	-	Continua
Abdel-Aty y Keller (2004)	0, 1, 2, 3, 4	-
Mitra y Washington (2007)	-	Continua
Vieira Gomes et al. (2012)	-	Discreta
Vieira Gomes (2013)	-	Discreta
Cunto et al. (2014)	-	Discreta
Chen y Xie (2016)	-	Continua
Tasic y Porter (2016)	-	Continua

El número de accidentes se identificó representado de dos maneras. Primero, como tasas de accidentes por intersección (un número continuo), o también como un número entero, asociado a la cantidad de accidentes por intersección. Kumara y Chin (2003) agregaron en su modelo en un

nivel todas aquellas intersecciones con seis o más accidentes. Con esto buscaron disminuir la sobre dispersión y el efecto de los valores extremos. En la Tabla 2.3 se muestra las variables independientes encontradas.

2.6 Variables influyentes en la ocurrencia de accidentes

Se encontraron tres grupos de variables que explican los accidentes. La infraestructura, que se refiere a condiciones de diseño geométrico, de equipamiento vial y de seguridad en la intersección; la operación de la intersección en donde características como el flujo vehicular, peatonal, las condiciones de visibilidad así como el tipo de control de la intersección son relevantes; y el entorno, que se refiere al uso de suelo en el que se enmarca la intersección y a la presencia de servicios que influyen en el nivel de conflictos y accidentabilidad de la intersección, como por ejemplo la presencia de servicios de educación, salud o paraderos de transporte público entre otros.

Con estas tres categorías de variables se realizó una caracterización de las variables que describen la infraestructura, la operación y el entorno de las intersecciones urbanas. Para ello se utilizó la nomenclatura descrita en la Tabla 2.4.

Tabla 2.5 Nomenclatura utilizada para definir las variables independientes

Abrev	Nombre	Abrev	Nombre	Abrev	Nombre	Abrev	Nombre
Acc	Acceso	Esq	Esquina	Lim	Límite	Salud	Instituciones de
Bike	Bicicleta	Estac	Estacionamiento	Med	Mediana	SC	Sin Control
Camin	Caminata	Exc	Exclusivo/a	N°	Número	Sem	Semaforizada
Can	Canalización	FPD	Flujo peatonal	P	Pistas	Sum	Suma
Categ	Categoría	Flu	Flujos	Peat	Peatonal	TMD	Tráfico medio diario
Conect	Conectividad	GD	Giro a la Derecha	Pop	Población	V	Velocidad
Dens	Densidad	GI	Giro a la Izquierda	Rec	Recorrido	Ver	Verde
Dist	Distancia	Grad	Gradiente	RP	Rama	Vere	Vereda
Educa	Insti. Educativas	Inter	Intersección	RS	Rama	Vis	Visibilidad

2.6.1 Variables asociadas a la infraestructura

La Tabla 2.5 muestra las variables relacionadas con la infraestructura que resultaron estadísticamente significativas en los distintos modelos analizados. La variable número de pistas resulta significativa en los modelos de Mitra y Washington (2007), quienes utilizan el número total de pistas por rama. En el trabajo de Kumara y Chin (2003), esta variable solo resultó significativa estadísticamente en la rama principal.

Para modelos como el propuesto por Abdel-Aty y Keller (2004), la presencia de mediana resultó significativa en la rama secundaria y principal, como un factor que disminuye la cantidad de accidentes. Esto se da para el caso de accidentes de mayores gravedades (medianamente graves, graves y fatales). En cambio para las otras dos categorías (sin lesionados y leves) no resultó significativa. Por lo general, en modelos como el de Vieira Gomes (2013) y Cunto *et al.* (2014), la presencia de mediana en la rama principal resultó significativa. Kumara y Chin (2003), Mitra y Washington (2007) y Vieira y Gomes (2013) utilizaron variables *dummy* para definir la presencia de la mediana y además, la variable ancho de mediana.

Las variables asociadas a pistas de viraje se encontraron representadas en modelos en distintos formatos, pero en todos ellos siempre resultó ser significativa. La presencia de pistas de viraje exclusivo a la izquierda disminuye considerablemente el número de accidentes en los modelos de Salifu (2003), Abdel-Aty y Keller (2004) y Kumara y Chin (2003). Del mismo modo, cuando el viraje a la derecha es canalizado, resultaba como un factor de disminución de accidentes. Esta variable la separaron en presencia de canalización de viraje a derecha en rama principal y secundaria. Cabe destacar que, en el modelo de AASHTO (2010), la variable es el número total de pistas que se utilizan para virar a la izquierda en la intersección y, análogamente, para el viraje a la derecha. El número de accesos también es una variable que influye en la accidentabilidad de una intersección, demostrando que si las configuraciones de las intersecciones son más complejas, entonces mayor es el número de accidentes esperados (Greibe, 2003).

Tabla 2.6 Variables de infraestructura utilizadas en modelos de accidentes

Variables Independientes	Greibe (2003)	Kumara y Chin (2003)	Salifu (2003)	Abdel-Aty y Keller (2004)	Mitra y Washington (2007)	Vieira Gomes et al. (2012)	Vieira Gomes (2013)	Cunto et al. (2014)	Tasic y Porter (2016)
% de vías expresa									•
Ancho de calle secundaria			•						
Ancho mediana en RP (Análogo RS)		•	•		•		•		
Ancho vereda en RP (Análogo RS)					•				
Ancho de P (m)							•		
Curva horizontal		•							
Dist. media de P viraje exclusiva izquierda		•							
Dist media de cruce peatonal							•		
Mediana en RP (Análogo RS)	•			•		•	•	•	
Facilidades para bicicletas	•								
Pista exc GD en RP (Análogo RS)		•		•					
Pista exc GI en RP (Análogo RS)		•	•						
Nº de Accesos	•						•		
Nº Pistas de viraje exclusivo	•								
Nº Pistas exc GI en RP (Análogo RS)				•					
Nº Pistas GI en RP (Análogo RS)				•				•	
Nº Pistas RP (Análogo RS)	•	•		•	•	•	•	•	
Pista de aceleración		•							

2.6.2 Variables asociadas a la operación

En la Tabla 2.6 se presentan las variables de operación encontradas en los distintos trabajos revisados. Éstas se relacionan directamente con la exposición. La más utilizada es el flujo vehicular.

Chen y Xie (2016) estudiaron la influencia del flujo vehicular en distintas versiones, para predecir accidentes en intersecciones. Utilizaron la variable de flujo como un flujo medio diario por intersección, por rama principal y secundaria, y también agregaron el cociente entre el flujo de la rama secundaria y el total, para así notar cómo influía el desequilibrio de los flujos entre las calles con el número de accidentes. Estas variables generaron un mejor ajuste en la modelación cuando le aplicaron el logaritmo del flujo, también demostrado en Tasic y Porter (2016).

En general, la velocidad de operación o velocidad límite no resultó significativa en los modelos. Los casos revisados donde tuvo significancia fueron los de Abdel-Aty y Keller (2004) y Salifu (2003). No obstante, resulta interesante según Salifu (2003) analizar una variable que considere la diferencia de velocidades entre cada rama, puesto que este diferencial puede generar conflictos de operación que causen accidentes.

La variable de visibilidad en modelos como el de Kumara y Chin (2003) no clasificaba dentro de aquellas de mayor importancia. Solo resultaba significativa considerando el 80% de confiabilidad. En cambio, Cunto *et al.* (2014) la utilizan como un factor modificador de accidentes por presencia de obstáculos, obteniendo una mejor significancia que Kumara y Chin (2003).

El tipo de control lo utilizan en algunos modelos (ver Tabla 2.6) como una variable *dummy*, es decir si la intersección es semaforizada o no (Mitra y Washington, 2007). Vieira Gomes (2013) utilizó esta variable asociando un número al tipo de control, considerando las intersecciones semaforizadas, con ceda el paso, pare y sin control. Por otro lado, Kumara y Chin (2003) usaron la variable asociada al número de fases dentro del ciclo del semáforo. En ambos casos estas variables resultaban significativas.

Tabla 2.7 Variables de operación significativas

Variables de operación	Greibe (2003)	Kumara y Chin (2003)	Salifu (2003)	Abdel-Aty y Keller (2004)	Mitra y Washington (2007)	Vieira Gomes et al. (2012)	Vieira Gomes (2013)	Cunto et al. (2014)	Tasic y Porter (2016)	Chen y Xie (2016)
Categ de la vía								•		
% viajes caminata									•	
% viajes tte. pub. trabajo									•	
Dist. rec. diaria vehículos									•	
Flujo peatonal diario total							•			
Giro derecha en RP/RS						•	•			
Giro izquierda en RP/RS						•	•			
Límite de velocidad			•	•						
N° de cruces peatonales							•			
N° giros izq. en intersecciones				•						
Cruce peatonal en RP (Análogo RS)							•			
TMDRS / TMDRP						•				•
TMDRS / TMDtot						•				
Tipo de control (Dummy)	•		•		•	•	•	•		
TMDmotor,tot + FPD,tot							•			
TMD veh.pesados			•							
TMD RP (Análogo RS)	•		•	•	•	•	•	•		•
TMD giro derecha		•								
TMD giro izquierda		•								
TMD motorizado	•	•	•			•				
TMD no motorizado	•									
TMD opuesto			•							
Visibilidad		•						•		

2.6.3 Variables asociadas al entorno

Las variables de entorno están presentes en el modelo de AASHTO (2010) y de Tasic y Porter (2016) y están directamente relacionadas con la ocurrencia de accidentes entre vehículos y peatones, pues se enfocan en las relaciones entre los distintos tipos de usuarios que interactúan en una intersección. En la Tabla 2.7 se muestran las variables significativas encontradas en los modelos revisados.

Tabla 2.8 Variables de entorno que explican la ocurrencia de accidentes en intersecciones

Variables de entorno	Greibe (2003)	Kumara y Chin (2003)	Salifu (2003)	Abdel-Aty y Keller (2004)	Mitra y Washington (2007)	Vieira Gomes <i>et al.</i> (2012)	Vieira Gomes (2013)	Cunto <i>et al.</i> (2014)	Tasic y Porter (2016)
Accidente ubicado por comuna (dummy)				•					
Densidad poblacional									•
Uso de suelo									•
N° de estacionamientos de bicicleta									•
N° de paraderos								•	•
N° de educa								•	
N° de centros de salud								•	
N° de locales de expendio de alcohol								•	
Presencia de estacionamientos								•	
Área total de veredas									•
Densidad intersecciones por km de calle									•

Estas variables de entorno son, el tipo de uso de suelo, la presencia de paraderos o de centros educacionales y/o de salud. Abdel-Aty y Keller (2004) utilizaron una variable *dummy*, asociada a la ubicación por comuna, considerando así el efecto geográfico del accidente. Cabe destacar que este tipo de variables se usa en menor medida, pues de los modelos analizados solo dos utilizan variables del entorno y uno utiliza la comuna como variable de localización.

La variable número de paraderos resultó significativa en ambos modelos que incluían variables del entorno. En cambio, las variables como el número de estacionamientos de bicicletas y de vehículos, no resultaron significativas en estos modelos. Por lo que ambos en ambos casos, se recomendó utilizar sólo variables de flujo, pues en aquellas variables de estacionamientos ya se considera implícitamente el efecto generado por el número de vehículos.

2.7 Conclusiones

Se identificaron tres grupos de modelos utilizados para el análisis de los accidentes en intersecciones urbanas: modelos tipo probit, de regresiones múltiples y finalmente, modelos lineales generalizados.

Se identificaron además tres clasificaciones de variables asociadas a las intersecciones aquellas relacionadas a la infraestructura, a la operación y al entorno.

Los coeficientes asociados a las variables revisadas tienen distintos resultados de significancia estadística dependiendo de las severidades de los accidentes (incluso cambiando el signo de su efecto en distintas severidades).

Los autores utilizan, en promedio, datos de cinco años. Con esto pueden estimar los modelos con una cantidad elevada de datos y a su vez, no generar sesgos debido a los problemas de utilizar muchos años de accidentes, ya que pueden existir cambios importantes en las características (variables) de las intersecciones analizadas.

Finalmente aquellas variables que resultaron relevantes y que son potencialmente utilizables en este trabajo de investigación son las de: presencia de pistas exclusivas de giro a la izquierda, presencia de mediana, número de pistas, flujo vehicular, flujo total de la intersección, tipo de control de la intersección, visibilidad y número de paraderos.

CAPÍTULO 3 DISEÑO DE UN EXPERIMENTO PARA ANALIZAR LOS ACCIDENTES

3.1 Introducción

En este capítulo se explica qué son cada una de las variables definidas y cómo medirlas y/o obtenerlas, así como el diseño experimental. Se presenta la definición formal de cada una de las variables agrupadas en las categorías infraestructura, operación y entorno de las intersecciones.

3.2 Definición de variables

En la Tabla 3.1 se muestra la variable dependiente, mientras que en la Tabla 3.2 se definen las variables de infraestructura, en la Tabla 3.3 las del entorno y en la Tabla 3.4 aquellas de operación.

Tabla 3.1 Definición de variable dependiente del experimento

Objeto de estudio	Variable explicativa	Periodo	Tipo de variable	Obtención
Intersección	Número de accidentes total	2009-2013	Discreta	BD Carabineros

Tabla 3.2 Definición de variables independientes de infraestructura

Variabes de infraestructura	Tipo de variables	Obtención
Número de ramas de acceso	Discreta	Diagramas de movimientos, mediciones
Número de pistas en la rama principal	Discreta	Inspección en terreno
Número de pistas en la rama secundaria	Discreta	Inspección en terreno
Número de calles unidireccionales	Discreta	Imágenes satelitales
Angulo de intersección	Discreta	Imágenes satelitales
Presencia de pistas exclusivas de viraje a izquierda en rama principal	Dicotómica	Inspección en terreno
Presencia de pistas exclusivas de viraje a izquierda en rama secundaria	Dicotómica	Inspección en terreno
Presencia de pistas exclusivas de viraje a derecha en rama principal	Dicotómica	Inspección en terreno
Presencia de pistas exclusivas de viraje a derecha en rama secundaria	Dicotómica	Inspección en terreno
Tipo de intersección	Categorica (3 niveles)	Imágenes satelitales
Presencia de valla peatonal	Dicotómica	Inspección en terreno
Presencia de mediana en rama principal	Dicotómica	Inspección en terreno
Presencia de mediana en rama secundaria	Dicotómica	Inspección en terreno

Tabla 3.3 Definición de variables independientes relacionadas al entorno

Variables del entorno	Tipo de variables	Obtención
Tipo de zona	Categoría (3 niveles)	Inspección en terreno
Número de paraderos	Discreta	Inspección en terreno
Número de centros de educación	Discreta	Inspección en terreno
Número de centros de salud	Discreta	Inspección en terreno

Tabla 3.4 Definición de variables independientes relacionadas a la operación

Variables de operación	Tipo de variables	Obtención
Presencia de giro a izquierda en rama principal	Dicotómica	Diagramas de movimientos, mediciones
Presencia de giro a izquierda en rama secundaria	Dicotómica	Diagramas de movimientos, mediciones
Número de cruces peatonales	Discreta	Inspección en terreno
Visibilidad (m)	Continua	Imágenes satelitales
Nivel de visión en intersección	Categoría (4 niveles)	Inspección en terreno
Nivel de densidad peatonal	Categoría (3 niveles)	Inspección en terreno
Flujo de la intersección	Continua	Mediciones vehiculares
Flujo que gira a derecha	Continua	Mediciones vehiculares
Flujo que gira a izquierda	Continua	Mediciones vehiculares
Flujo de la rama principal	Continua	Mediciones vehiculares
Flujo de la rama secundaria	Continua	Mediciones vehiculares
Flujo de vehículos livianos	Continua	Mediciones vehiculares
Flujo de camiones	Continua	Mediciones vehiculares
Flujos de taxi colectivos	Continua	Mediciones vehiculares
Flujo de taxi buses	Continua	Mediciones vehiculares
Flujo de buses	Continua	Mediciones vehiculares
Flujo no motorizado	Continua	Mediciones vehiculares
Flujo de la rama secundaria / Flujo de la rama principal	Continua	Mediciones vehiculares
Flujo de la rama secundaria / Flujo de la intersección	Continua	Mediciones vehiculares

3.3 Matriz factorial

El diseño factorial completo es un método experimental que consiste en cruzar todos los niveles de todas las variables explicativas, con el fin de estudiar la interacción de cada una de ellas. Se realizó un diseño factorial completo, identificando dos grupos de variables independientes: las variables relacionadas al entorno en conjunto con las variables de infraestructura y las variables operativas,

cada una de ellas con diferentes niveles. El diseño experimental busca que exista variabilidad en los atributos para asegurar la validez del modelo. El resultado de esto permite conocer la cantidad mínima de datos que se requieren en cada celda de la matriz, es decir, en cada combinación factible de estos atributos.

La matriz factorial tiene en las columnas la interacción de todas las variables relacionadas a la infraestructura en conjunto con aquellas del entorno. En las filas se agregan las variables relacionadas a la operación interactuando entre sí. En el Anexo A.3.1 se presenta la matriz factorial conceptual y la matriz factorial completa se presenta en el Anexo Magnético, Capítulo 3.

Para la elaboración de la matriz, se definieron todas las variables continuas con tres niveles usando el algoritmo de k-medias (Hartigan y Wong, 1979). Las demás variables presentaron los niveles con las cuales se definieron y las dicotómicas con sus respectivos dos niveles.

La matriz obtenida tiene 688,747,536 filas y 2,519,424 columnas considerando un diseño factorial con interacciones entre todas las variables mencionadas. Esto genera una matriz de $1.735 \cdot 10^{15}$ celdas. El cálculo teórico del tamaño muestral se determinó a través del análisis de potencia empleando los métodos de Cohen (1988), resultando dos réplicas por celda al distribuir homogéneamente en la matriz las muestras, es decir, se deben tener $3.47 \cdot 10^{15}$ datos. Por otro lado, es importante recalcar que este resultado es teórico y en la práctica los modelos pueden ser validados y pueden entregar patrones y estimaciones adecuadas con una menor cantidad de datos.

Ahora bien, debido a la elevada cantidad de variables en la matriz, ésta resulta con una elevada cantidad de celdas. No obstante existe una cantidad importante de combinaciones (configuraciones de las intersecciones) que no resultan factibles de existir, como por ejemplo: intersecciones del tipo T con más de tres ramas de acceso, intersecciones con las dos calles unidireccionales y con presencia de flujo opuesto o, también, intersecciones sin presencia de giros a izquierda o derecha y con presencia de pistas exclusivas de viraje a izquierda o derecha respectivamente.

3.4 Conclusiones

En este capítulo se definieron las variables independientes del estudio, las que se clasificaron en 13 variables de infraestructura, 19 variables de operación y cuatro de entorno. Se definió la matriz factorial teórica para el set de variables definidas la cual obtuvo un tamaño de $1.735 \cdot 10^{15}$ celdas. Se desarrolló el cálculo del tamaño muestral mínimo, en el cual se obtiene una cantidad mínima de dos réplicas por celda. La cantidad de datos necesarias para cumplir con el diseño factorial es elevada, adquiriendo mayor importancia el realizar un análisis previo con el objetivo de descartar variables disminuyendo el número de éstas de tal modo de poder generar un experimento factible y teóricamente correcto con la cantidad de datos que se tiene.

Es importante comentar que el ideal es tener variabilidad en todos los atributos. Es por esto que el realizar el diseño experimental y el cálculo de la matriz factorial permite obtener una idea de cuántos datos y cuántas réplicas por celda son necesarias para asegurar estadísticamente la variabilidad, es decir, asegurar la validez del modelo. En este caso, es fundamental reducir el número de variables y realizar un análisis de todas las combinaciones que no son factibles de tal modo de obtener una nueva matriz factorial de menor tamaño y poder aplicar en la práctica este modelo. Esto debido a que la cantidad de datos disponible es muchos órdenes de magnitud menor a aquella que teóricamente resultaría necesaria.

CAPÍTULO 4 ELABORACIÓN DE LA BASE DE DATOS DE ACCIDENTES

4.1 Introducción

El objetivo de este capítulo es presentar la base de datos generada para el análisis de los accidentes. También se realiza una caracterización de los datos, tanto de accidentes como de aquellos relativos a las intersecciones, y se explica la composición de la base de datos construida.

4.2 Elaboración de la base de datos

Las componentes que estructuran la base de datos se presentan en la Figura 4.1.

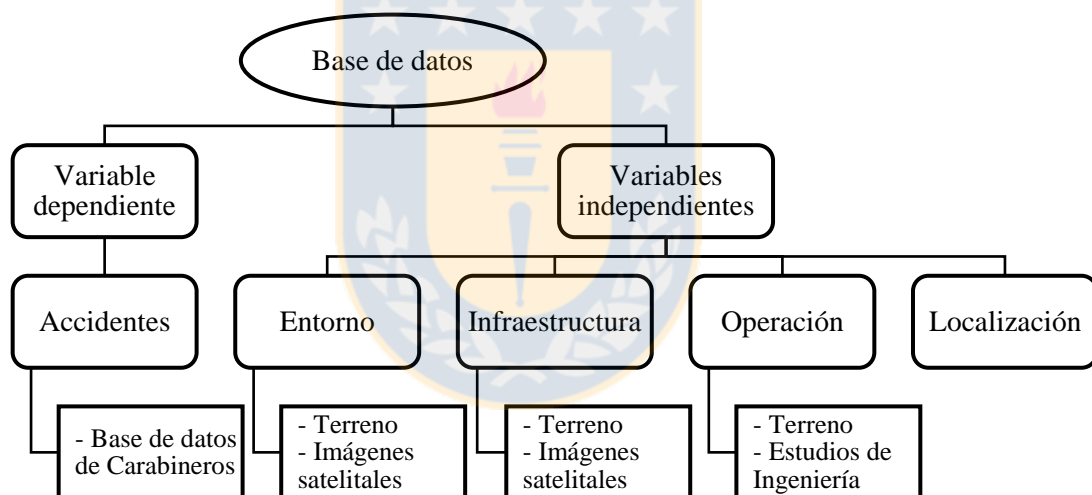


Figura 4.1 Estructura de la base de datos completa

La base de datos de accidentes fue obtenida de Carabineros de Chile y comprende los accidentes ocurridos en intersecciones del Gran Concepción entre los años 2009 y 2013. Por otro lado se construyó, en base a la localización de los accidentes y por consiguiente de las intersecciones, la base de datos de las variables independientes. Estas variables se obtuvieron de distintas fuentes (terreno, imágenes satelitales y estudios de ingeniería de SECTRA) clasificándose en tres grupos, entorno, infraestructura y operación.

Los accidentes fueron georreferenciados en el proyecto GESITRAN (2014). No obstante, se tuvo que depurar estos datos para considerar sólo aquellos que tenían estricta relación con las intersecciones urbanas definidas. Para ello se consideró un radio de 50 m en torno al centro de la intersección. Además, en el proceso de identificación de los accidentes se tuvo que sistematizar los nombres de las calles y así obtener una sola nomenclatura asociada a cada intersección.

Se analizó separadamente la influencia del periodo del día en los accidentes, configurando seis bases de datos: (1) periodo punta mañana, (2) periodo punta medio día, (3) periodo punta tarde, (4) periodo fuera de punta, (5) periodo diurno, y (6) día completo. La obtención del horario de estos periodos se realizó en base a la periodización de tres intersecciones con mediciones continuas (ver Anexo A.4.1). Los periodos considerados en cada base de datos se muestran en la Tabla 4.1.

Tabla 4.1 Bases de datos del estudio.

Base de datos	Periodo	Inicio	Fin
(1)	Punta mañana (PM)	7:15	9:15
(2)	Punta medio día (PMD)	12:30	14:30
(3)	Punta tarde (PT)	17:45	19:45
(4)	Fuera de punta (FP)	Resto del día	
(5)	Diurno	7:00	23:00
(6)	Completo	0:00	23:59

Ahora bien, la base de datos asociada a las características de las intersecciones se generó en terreno. Dependiendo de cada variable, debió realizarse un trabajo en terreno y/o con la ayuda de Google *Earth*, de tal modo de obtener los valores de las variables asociadas a infraestructura, operación y entorno. Además fue necesaria la información de catastros de semáforos de GESITRAN para obtener el número de fases, así como entender el funcionamiento y ordenamiento de los flujos en cada intersección. Del mismo modo, se utilizaron mediciones de SECTRA, tanto periódicas como continuas.

Esquemáticamente, la base de datos tiene en cada fila: el nombre de la intersección, el número de accidentes totales por año (2009, 2010, 2011, 2012 y 2013), el número de accidentes total en el periodo y luego tiene en las siguientes columnas todos los valores de las variables definidas de infraestructura y del entorno que son idénticas en los seis periodos (PM, PMD, PT, FP, Diurno y Día completa). Finalmente, están presentes las variables de operación calculadas para cada una de

las intersecciones en cada periodo. La base de datos tiene 273 filas, una por cada intersección en estudio y existe una base de datos para cada periodo definido. La base de datos resultante se encuentra en Anexo Magnético, Capítulo 4.

4.3 Definición de la variable dependiente

La variable dependiente es el número de accidentes de todos los tipos y severidades de cada intersección desde el año 2009 al 2013. Para la obtención de este valor fue necesario re-estructurar y depurar la base de datos de accidentes obtenida de Carabineros de Chile. Luego, se realizó el conteo y se obtuvo el valor asociado al número total de accidentes (todo tipo de accidente) en cada intersección, ocurrido durante el periodo de análisis en los cinco años.

4.4 Definición de las variables independientes

Las variables independientes se obtuvieron de mediciones, de imágenes satelitales y de la obtención de información de la plataforma GESITRAN. Se presentan a continuación las variables independientes de entorno, infraestructura y operación.

4.4.1 Variables asociadas al entorno de la intersección

En la Tabla 4.2 se presentan las variables del entorno de las intersecciones definidas.

Tabla 4.2 Variables del entorno de las intersecciones obtenidas para la base de datos

Nombre de Variable	Valor	Etiqueta
Tipo de Zona	1	Zona principalmente del tipo residencial
	2	Zona principalmente del tipo comercial
	3	Zona mixta (residencial y comercial)
Número de paraderos	0 - 3	Número de paraderos en radio de 50 m en torno al centro de la intersección
Número de centros de educación	0 - 3	Número de centros de educación en radio de 130 m en torno al centro de la intersección
Número de centros de salud	0 - 1	Número de centros de salud en radio de 130 m en torno al centro de la intersección

En el tipo de zona no se utilizó la categoría de zona industrial, pues con respecto a la base de datos de intersecciones que se disponían no había ninguna inserta en las zonas industriales del Gran Concepción. Para definir si la zona se consideraba mixta, residencial o comercial, se revisó el uso de suelo en torno a cada intersección, generando en un área de un radio de 150 m alrededor de ésta, una proporción entre zona residencial y comercial. Cuando ésta se encontraba entre 40% y 60% se consideró una zona mixta.

4.4.2 Variables asociadas a la infraestructura de la intersección

En la Tabla 4.3 se presentan las variables de infraestructura definidas.

Tabla 4.3 Variables de infraestructura de las intersecciones obtenidas para la base de datos

Nombre de Variable	Valor	Etiqueta
Tipo de Intersección	1	Intersección en T
	2	Intersección en X
	3	Intersección peatonal
Pista exclusiva de giro a derecha en rama principal	0	No existe
	1	Existe al menos una pista
Pista exclusiva de giro a derecha en rama secundaria	0	No existe
	1	Existe al menos una pista
Pista exclusiva de giro a izquierda en rama principal	0	No existe
	1	Existe al menos una pista
Pista exclusiva de giro a izquierda en rama secundaria	0	No existe
	1	Existe al menos una pista
Número de ramas de acceso	1 - 4	Número de accesos a intersección
Número de pistas de la rama principal	2 - 8	Número total de pistas en ambos sentidos
Número de pistas de la rama secundaria	0 - 6	Número total de pistas en ambos sentidos
Mediana en la rama principal	0	No existe o tiene menos de 0.8 m de ancho
	1	Existe y tiene al menos 0.8 m de ancho
Mediana en la rama secundaria	0	No existe o tiene menos de 0.8 m de ancho
	1	Existe y tiene al menos 0.8 m de ancho
Valla peatonal	0	No existe
	1	Existe al menos en alguna acera
Ángulo de la intersección	0° - 90°	Ángulo formado entre ambos ejes de las calles que conforman la intersección
Número de calles unidireccionales	0 - 2	Número de calles de la intersección con sentido único

4.4.3 Variables asociadas a la operación de la intersección

En la Tabla 4.4 se presentan las variables de operación de las intersecciones definidas. Es necesario destacar que las mediciones de condiciones de operación fueron realizadas en el año de estudio, es decir durante el 2015. Por lo tanto, se asumió que las diferencias relativas de las variables de operación (densidad peatonal, flujos vehiculares, entre otros) entre las intersecciones del estudio se mantienen constantes.

Tabla 4.4 Variables de la operación de las intersecciones obtenidas para la base de datos

Nombre de Variable	Valor	Etiqueta
Giro a la izquierda en la rama principal	0	No existen giros a izquierda
	1	Existe al menos un giro a izquierda
Giro a la izquierda en la rama secundaria	0	No existen giros a izquierda
	1	Existe al menos un giro a izquierda
Densidad peatonal	1	Densidad peatonal baja en la intersección
	2	Densidad peatonal media en la intersección
	3	Densidad peatonal alta en la intersección
Número de cruces peatonales	0 - 8	Número de cruces peatonales sin considerar las medianas
	0 - 11	Número de cruces peatonales considerando las medianas
Número de fases del semáforo	2 - 3	Número de fases del semáforo en la intersección
Visibilidad (m)	4.8 - 200	Distancia de visibilidad, según método MOP (2015)
Nivel de visión	1	Nivel de visión malo, obstáculos importantes
	2	Nivel de visión regular, obstáculos menores
	3	Nivel de visión bueno, obstáculos casi imperceptibles
	4	Nivel de visión excelente, sin obstáculos
Flujo de la intersección (veq)	Continuo	Flujo total en la intersección
Flujo de la rama principal (veq)	Continuo	Flujo total de la rama principal
Flujo de la rama secundaria (veq)	Continuo	Flujo total de la rama secundaria
Flujo que gira a izquierda (veq)	Continuo	Flujo total que gira a izquierda en la intersección
Flujo que gira a derecha (veq)	Continuo	Flujo total que gira a derecha en la intersección
Flujo opuesto (veq)	Continuo	Flujo total que se opone en la intersección
Flujo de vehículos livianos (veh)	Continuo	Flujo total de vehículos livianos en la intersección
Flujo de taxi colectivos (veh)	Continuo	Flujo total de taxi colectivos en la intersección
Flujo de taxi buses (veh)	Continuo	Flujo total de taxi buses en la intersección
Flujo de camiones (veh)	Continuo	Flujo total de camiones en la intersección
Flujo de buses inter urbanos (veh)	Continuo	Flujo total de buses inter urbanos en la intersección
Flujo no motorizado (veh)	Continuo	Flujo total no motorizado en la intersección

La densidad peatonal se midió de manera visual en terreno de acuerdo a las categorías baja, media y alta según MINVU (2009). Esta medición se realizó para todas las intersecciones durante 10 min,

entre las 12:30 hrs y las 15:00 hrs y sólo en días laborales. El número de cruces peatonales en cada intersección se consideró de dos maneras. (1) La primera fue considerar un cruce peatonal a lo ancho de cada calzada sin considerar las medianas. (2) La segunda manera fue considerar la separación de un cruce peatonal en dos siempre y cuando existiese algún tipo de refugio para el peatón. Para obtener estas variables fue necesario revisar las configuraciones de todas las intersecciones en terreno, de modo tal de observar el funcionamiento y comportamiento de los peatones en ellas. La visibilidad corresponde a la menor distancia en la intersección, según MOP (2015). Para el cálculo de la distancia de visibilidad se considera a 3 m desde la línea de parada el punto de inicio y se proyecta una línea desde este punto pasando por el obstáculo que exista hasta el centro de la pista por la que transita el flujo principal. Luego, el cateto opuesto al triángulo formado corresponde a la distancia de visibilidad (d_1 y d_2); ver Figura 4.2. Se asignó como valor de distancia de visibilidad de la intersección el valor del cateto opuesto más pequeño. Esta variable se midió desde Google Earth.



Figura 4.2 Distancia de visibilidad, MOP (2015)

4.5 Obtención de los flujos vehiculares

Los flujos vehiculares se obtuvieron de mediciones de SECTRA (2014 y 2016). Los datos se actualizaron y homologaron todos al año 2015. Para ello se utilizó la información de SECTRA (2016) con las tasas de crecimientos vehiculares de la Tabla 4.5.

Tabla 4.5 Tasas de crecimiento vehicular adoptadas en este estudio (SECTRA, 2016)

Tipo de vehículo	Tasa de crecimiento (%)
Livianos	3.5
Camiones de dos ejes	4.9
Camiones de más de dos ejes	5.1
Transporte Público	0

En la utilización de esta variable existe un supuesto, donde se asume que las diferencias relativas de esta variable en las intersecciones se han mantenido a lo largo de los años que se están considerando en el análisis (2009 a 2013).

Los flujos de años distintos al 2015 se expandieron, tanto positiva como negativamente, usando las tasas de crecimientos anuales de la Tabla 4.5. De las mediciones se calculó el valor representativo de cada periodo en estudio así como también el del día completo. Debe notarse que en este trabajo se utilizaron solo mediciones de días laborales. En el caso de las mediciones continuas, simplemente se sumó el flujo respectivo abarcando cada rango horario de los periodos analizados (punta mañana, punta medio día, punta tarde, fuera de punta, diurno y día completo).

Los rangos horarios que se escogieron para definir cada periodo de estudio se obtuvieron de la periodización de tres intersecciones realizada según el método de la hora móvil de mayor flujo equivalente, cuyos resultados se encuentran en el Anexo A.4.1. Se debe recordar que este proceso se realizó para elaborar las distintas bases de datos presentadas anteriormente (ver sección 4.2). Por lo tanto, se tiene para cada intersección el valor del flujo correspondiente y esto se calculó para cada una de las seis bases de datos (correspondientes a cada periodo analizado).

4.6 Caracterización de las intersecciones

Las intersecciones en las cuales se dispuso y pudo obtener la información necesaria para el análisis, y por consiguiente, las intersecciones incluidas en la base de datos, se distribuyen en las comunas del Gran Concepción como se muestra en la Tabla 4.6.

Tabla 4.6 Cantidad de intersecciones por comuna y tipo en base de datos

Comuna	N° de Intersecciones	Tipo		
		T	Cruz	Peatonal
Concepción	177	41	132	4
Chiguayante	24	6	16	2
San Pedro	19	14	5	0
Tomé	6	0	6	0
Talcahuano	46	27	13	6
Hualqui	1	0	1	0
Total	273	88	173	12

Además, se realizó un análisis espacial por comuna, en donde se presentan todos los accidentes de las intersecciones de la muestra distribuidos en la red vial; las imágenes se encuentran en Anexo A.4.2. De este análisis espacial, se observan los ejes con mayor nivel de accidentabilidad y las intersecciones con mayores niveles de accidentes para cada eje. En este sentido, en la comuna de Concepción, Los Carrera, Paicaví y Prat concentran altos niveles de accidentes. Existen también intersecciones, como 21 de Mayo con Pedro de Oña o Collao con Tegalda, en donde se concentran gran cantidad de eventos. En la comuna de San Pedro de la Paz, en el eje Pedro Aguirre Cerda existen dos intersecciones con altos niveles de accidentes: Daniel Belmar y Diagonal Bío Bío. En Talcahuano, la mayor cantidad de accidentes se concentra en la intersección de Colón con Las Hortensias. En Chiguayante la intersección que presenta mayor cantidad de accidentes es Santa Elena con Manuel Rodríguez.

4.7 Caracterización de accidentes

Los accidentes se caracterizaron según su evolución temporal por año, su distribución espacial y su severidad. El objetivo de realizar esta caracterización radica en la necesidad de comparar el comportamiento de la base de datos de accidentes del estudio con respecto a todos los datos de accidentes en intersecciones urbanas del Gran Concepción. De aquí en adelante se utiliza el término “muestra” para referirse a la base de datos sintetizada en el presente trabajo. Esto debido a que los accidentes con los que se cuenta son solo aquellos que ocurrieron en 273 intersecciones del Gran Concepción, a las cuales se calculó y obtuvo todas las variables definidas en la sección 4.4.

4.7.1 Análisis temporal de accidentes

Un aspecto importante es la representatividad o no de la muestra con respecto a los accidentes totales del Gran Concepción en intersecciones urbanas. En este caso se busca realizar la comparación temporal entre las cantidades de accidentes anuales en intersecciones urbanas, como se presenta en la Tabla 4.7. Se observa una proporción que oscila entre el 53.3% y 39.0% de

representatividad por año entre la base de datos de accidentes en intersecciones urbanas del Gran Concepción y la base de datos de accidentes de la muestra.

Tabla 4.7 Cantidad de accidentes por año de la población y de la muestra

Año	2009	2010	2011	2012	2013	Total
BD Gran Concepción	1457	1421	1283	1439	1839	7439
Muestra	777	661	524	561	908	3431
Porcentaje	53.3%	46.5%	40.8%	39.0%	49.4%	46.1%

Al comparar la evolución temporal de manera gráfica (ver Figura 4.3), se observa que las curvas resultan prácticamente iguales en sus pendientes y formas, luego el comportamiento de los datos de la muestra con respecto al total de accidentes año a año durante el periodo 2009-2013 resulta representativo. Desde un punto de vista estadístico se podría aplicar el *Log-Rank Test P* y determinar si existen diferencias significativas entre ambas curvas.

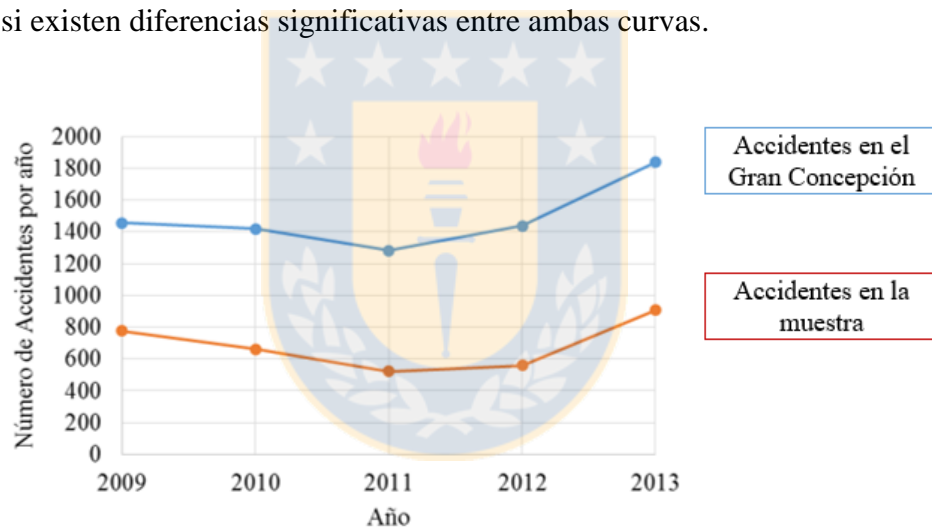


Figura 4.3 Evolución temporal comparativa de accidentes en intersecciones urbanas

4.7.2 Análisis espacial de accidentes

Análogamente, se realiza la comparación espacial de los accidentes en intersecciones urbanas. Para ello se presenta en la Tabla 4.8 los números de accidentes por comuna captados en la base de datos en comparación con los datos totales de accidentes en intersecciones urbanas.

Tabla 4.8 Accidentes por comunas en la BD del Gran Concepción y en la muestra

Comuna	BD Gran Concepción	Muestra	Porcentaje (%)
Concepción	3710	2385	64.3
Chiguayante	428	220	51.5
San Pedro de la Paz	1300	411	31.6
Tomé	541	67	12.4
Talcahuano	1460	348	23.8
Hualqui	0	0	0.0
Total	7439	3431	46.1

La comuna que tiene la mayor cantidad de accidentes es Concepción, lo que es esperable, pues es en donde existe mayor proporción de flujos vehiculares. Además, en la comuna se concentra la mayor cantidad de servicios, puestos de trabajo y centros de educación de la zona. En la muestra, Concepción tiene el 64% de los accidentes totales, en cambio en Tomé o Talcahuano, la muestra solo capta el 12 y 24% respectivamente. Finalmente, se tiene un total de 46% de accidentes captados en la muestra. La comuna de Hualqui no tiene accidentes reportados en los años de estudio.

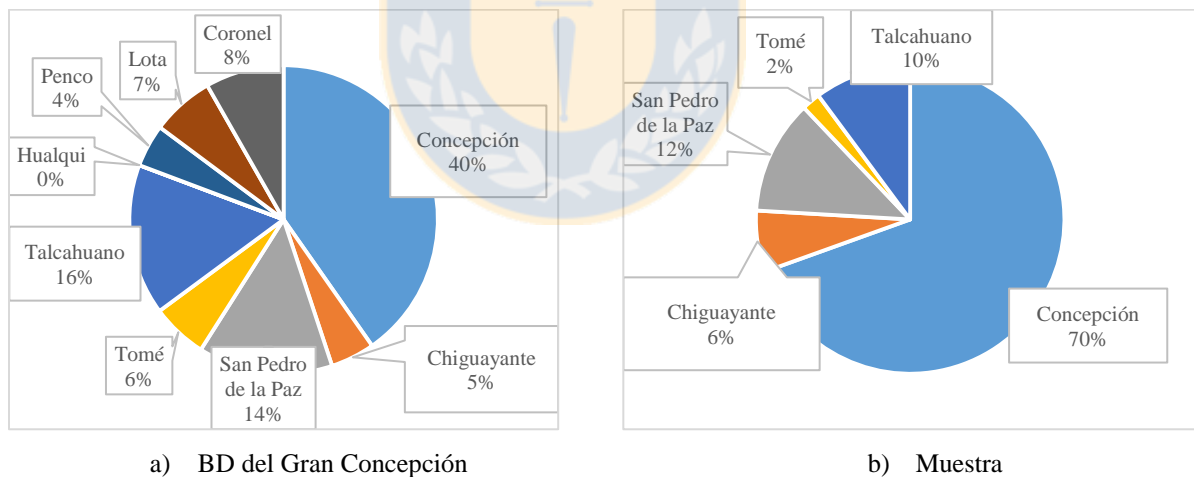


Figura 4.4 Accidentes en intersecciones urbanas por comuna en el periodo de estudio

En las Figuras 4.4.a) y 4.4.b) se observa que las proporciones de accidentes se mantienen casi constantes en las comunas de Chiguayante, Talcahuano y San Pedro de la Paz. No obstante, en el caso de Concepción existe diferencia en el porcentaje de accidentes, el cual es mayor en la muestra que el de la base de datos de accidentes completa. En consecuencia, existe una tendencia en este análisis a primar aquellos accidentes en dicha comuna, por sobre los demás. Luego, en lo que

respecta a espacialidad, no se puede incluir ni realizar un correcto análisis ya que los datos están sesgados.

4.7.3 Análisis de la severidad de los accidentes

Análogamente, en lo que respecta la severidad de los accidentes, se busca observar el comportamiento de estos en la muestra con respecto a la base de datos total del Gran Concepción.

En la Figura 4.5 se presentan las gráficas de las cantidades de accidentes según cada clase de severidad, tanto para la base de datos total del Gran Concepción, como para la base de datos en estudio.

Como se observa, el 50% de los accidentes son leves, mientras que un 30% son accidentes sin lesionados. La cantidad de accidentes disminuye a medida que aumenta la severidad, pasando a los graves no incapacitantes con un 7%, graves con un 11% y fatales con 2%. En el caso de la muestra, los porcentajes prácticamente no varían y se mantienen, lo cual nuevamente demuestra la representatividad de la muestra con respecto a los datos totales.

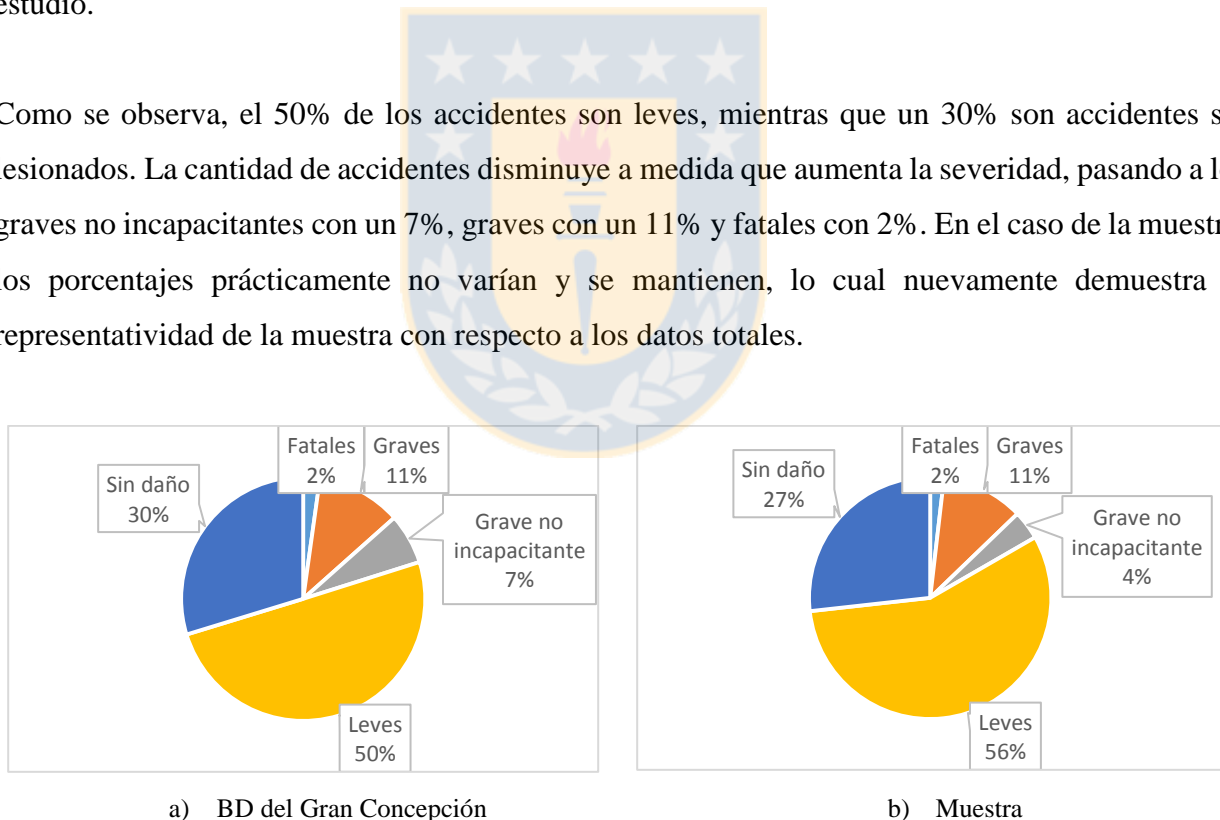


Figura 4.5 Accidentes en intersecciones urbanas según severidad

La Figura 4.6 muestra la cantidad de accidentes en la muestra según la severidad y una línea recta que representa la tendencia lineal de los datos.

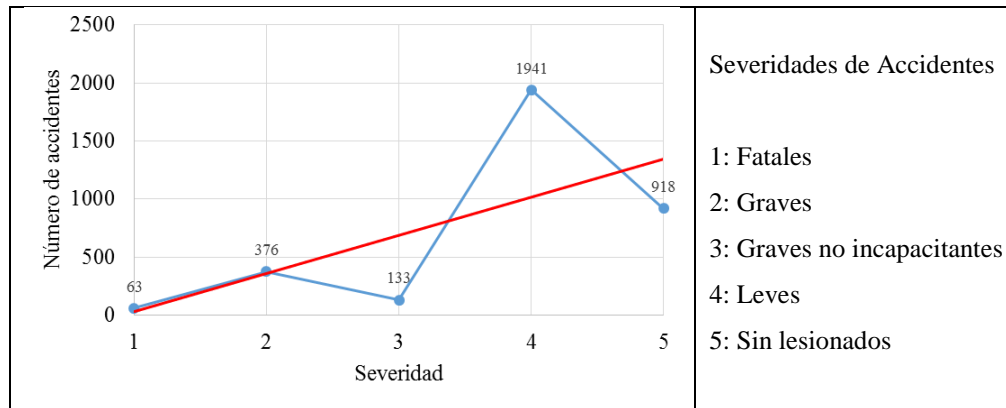


Figura 4.6 Accidentes de la base de datos de la muestra según su severidad

Se observa un comportamiento fuera de la tendencia, puesto que se espera un aumento sostenido en la cantidad de accidentes según disminuye la severidad. Existe un sesgo en los datos en lo que respecta la cantidad de accidentes en la categoría graves no incapacitantes, pues hay muy pocos accidentes reportados según esta categoría. En cambio, existe un elevado número de accidentes leves. El motivo de estos valores puede deberse a la toma del dato de accidente. Esto, porque su definición en la categoría como accidente “medianamente grave” (en planilla de accidentes de Carabineros de Chile) genera que quien reporte el dato, escoja entre la categoría leve, medianamente grave o grave, induciendo entonces errores según el criterio de cada persona que defina la severidad del accidentes. Este error ciertamente está inducido por un problema en los nombres asociados a las severidades medias de accidentes.

Por consiguiente, las categorías de accidentes leves tienen quizás más accidentes de lo que en verdad debiesen tener, perdiéndose accidentes en la categoría graves no incapacitantes. Finalmente, se puede concluir de esta caracterización que los datos de accidentes se pueden trabajar solo de manera agregada según las severidades en lo que respecta el Gran Concepción.

4.7.4 Análisis según tipo de intersección de accidentes

En la Tabla 4.9 se muestran los órdenes de magnitud de accidentes de cada base de datos, desagregando por tipo de intersección.

Tabla 4.9 Cantidad de accidentes por tipo de intersección en las seis bases de datos

Tipo de Intersección (N°)	PM	PMD	PT	FP	Diurno	Día Completo
T (88)	154	135	164	563	1004	1102
Cruz (173)	231	317	346	1164	2047	2244
Peatonal (12)	9	17	23	34	83	85

Se tiene un número elevado de accidentes para las intersecciones en T y aquellas en cruz, esto está relacionado con el número de intersecciones que se dispone, en donde las intersecciones en cruz son 173 y las en T son 88. Por ello, el menor número de accidentes es en intersecciones peatonales. Además como se presenta en la Tabla 4.10 los promedios de accidentes por tipo de intersección indican que aquellas en cruz es en donde más accidentes en promedio ocurren seguido de las intersecciones en T.

Tabla 4.10 Promedio de accidentes por tipo de intersección en las seis bases de datos

Tipo de Intersección	PM	PMD	PT	FP	Diurno	Día Completo
T	1.75	1.53	1.86	6.39	11.40	12.52
Cruz	1.33	1.83	2.00	6.72	11.83	12.97
Peatonal	0.75	1.41	1.91	2.83	6.91	7.08

4.8 Conclusiones

Se elaboraron seis bases de datos con información de distintas fuentes (Accidentes de Carabineros de Chile, mediciones de estudios mandados por SECTRA y de información recopilada en terreno). La diferencia entre cada base de datos generada radicó en el rango horario considerado, buscando en el análisis posterior, revisar y comparar las diferencias que puedan existir. Además se realizó una caracterización de los accidentes pertenecientes a las 273 intersecciones, los cuales representan una muestra equivalente al 46,1% de la totalidad de accidentes en intersecciones urbanas del Gran Concepción considerando el periodo desde el año 2009 al 2013.

Para construir la base de datos y cumplir con el diseño experimental propuesto, se tuvo que realizar una re-estructuración de los datos de accidentes de Carabineros de Chile, puesto que no existía una única nomenclatura de los nombres de las calles, lo cual dificultó la localización de los accidentes. También es importante notar que el resultado de tener 273 intersecciones está sujeto a la obtención de información del flujo vehicular. Además para los análisis posteriores no se pueden comparar y

analizar intersecciones si no tienen la misma cantidad de atributos o el mismo tipo de información asociada, esto pues se estaría generando un sesgo en aquellas variables que tendrían asociadas una menor cantidad de valores.

Del análisis de severidad de los accidentes en intersecciones urbanas se observó que existe un número reducido de accidentes fuera de la tendencia en la categoría “lesiones medianamente graves”, debido a una posible sub estimación de éstos al ser reportados como accidentes con lesionados leves y/o como accidentes con lesionados graves.



CAPÍTULO 5 ANÁLISIS ESTADÍSTICO

5.1 Introducción

En este capítulo se muestran los resultados del análisis exploratorio y estadístico de las variables de la base de datos. Se presentan los resultados del análisis de valores extremos, del análisis de los efectos principales de cada variable sobre la media y, finalmente, se presentaron los resultados del análisis bi-variado entre las variables explicativas.

5.2 Procedimiento de análisis

La Figura 5.1 muestra el procedimiento utilizado en el análisis que se presenta en este capítulo.

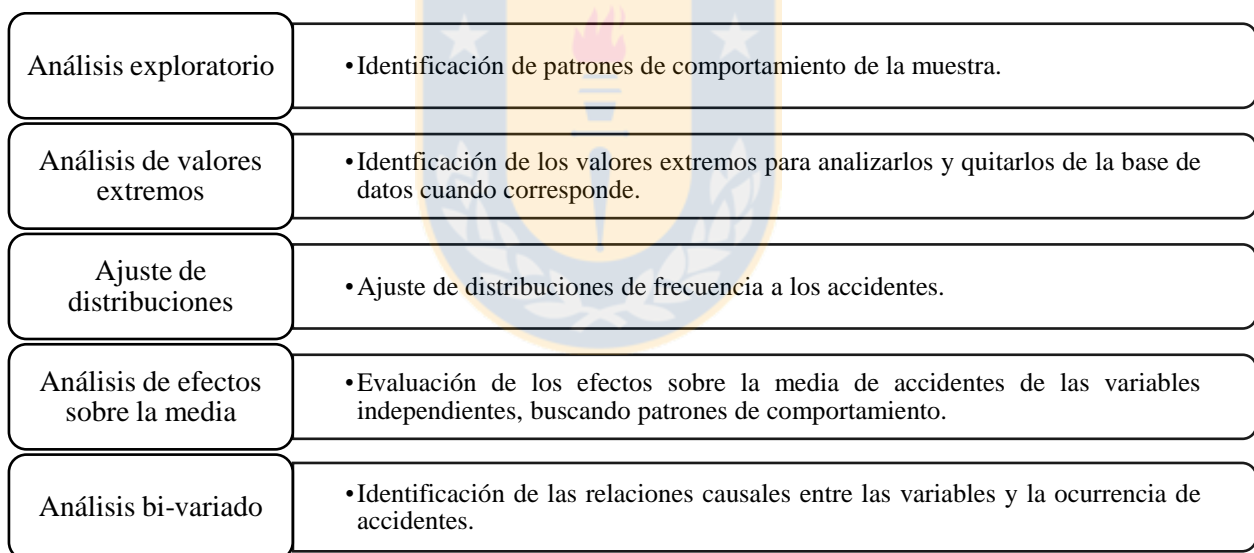


Figura 5.1 Esquema de la metodología del análisis de las variables del experimento

El análisis exploratorio consistió en obtener los histogramas de frecuencias de todas las variables y obtener información sobre qué proporciones de intersecciones se tienen con las características de cada variable definida. En segunda instancia, el análisis de valores extremos buscó realizar un filtro de intersecciones que presentaran un número muy elevado de accidentes con respecto a las demás, buscando mejorar los ajustes de los análisis posteriores. Con la base de datos depurada, se procedió

a ajustar las distribuciones a las variables dependientes, buscando la mejor distribución, en base a los parámetros de los test de Kolmogorov-Smirnov y de Anderson-Darling. Luego, se realizó un análisis gráfico de los efectos sobre la media de accidentes de todas las variables independientes.

Finalmente, se realizó un análisis bi-variado buscando obtener las relaciones causales de las variables independientes con los accidentes en cada periodo. Para el análisis de las variables continuas se utilizó el estudio de las correlaciones según el test de Pearson. Para el análisis de las variables de dos niveles se utilizó la prueba t-Student. Esta prueba se sustenta en los supuestos de normalidad, de homocedasticidad, y en la independencia de las muestras.

En última instancia, para el análisis de las variables con más de dos niveles se utilizó el ANOVA. Éste se sustenta en la homogeneidad de varianzas entre las variables si esto no se cumple se utilizan pruebas robustas de igualdad de medias para revisar la significancia estadística de la variable. Sumado a esto se realiza un análisis de pruebas post hoc de Bonferroni o Games Howel, que busca verificar además las relaciones que existen entre los niveles propios de cada variable.

5.3 Estadística descriptiva de las variables dependientes

La Figura 5.2 muestra los resultados del análisis exploratorio de la variable dependiente (número de accidentes por intersección). En punta mañana, punta medio día y punta tarde los histogramas presentan gran cantidad de datos cercanos al valor cero, mientras que en los periodos fuera de punta, diurno y día completo, la forma de la distribución cambia, pues existe menor cantidad de ceros, lo que se debe a las horas consideradas en cada base de datos.

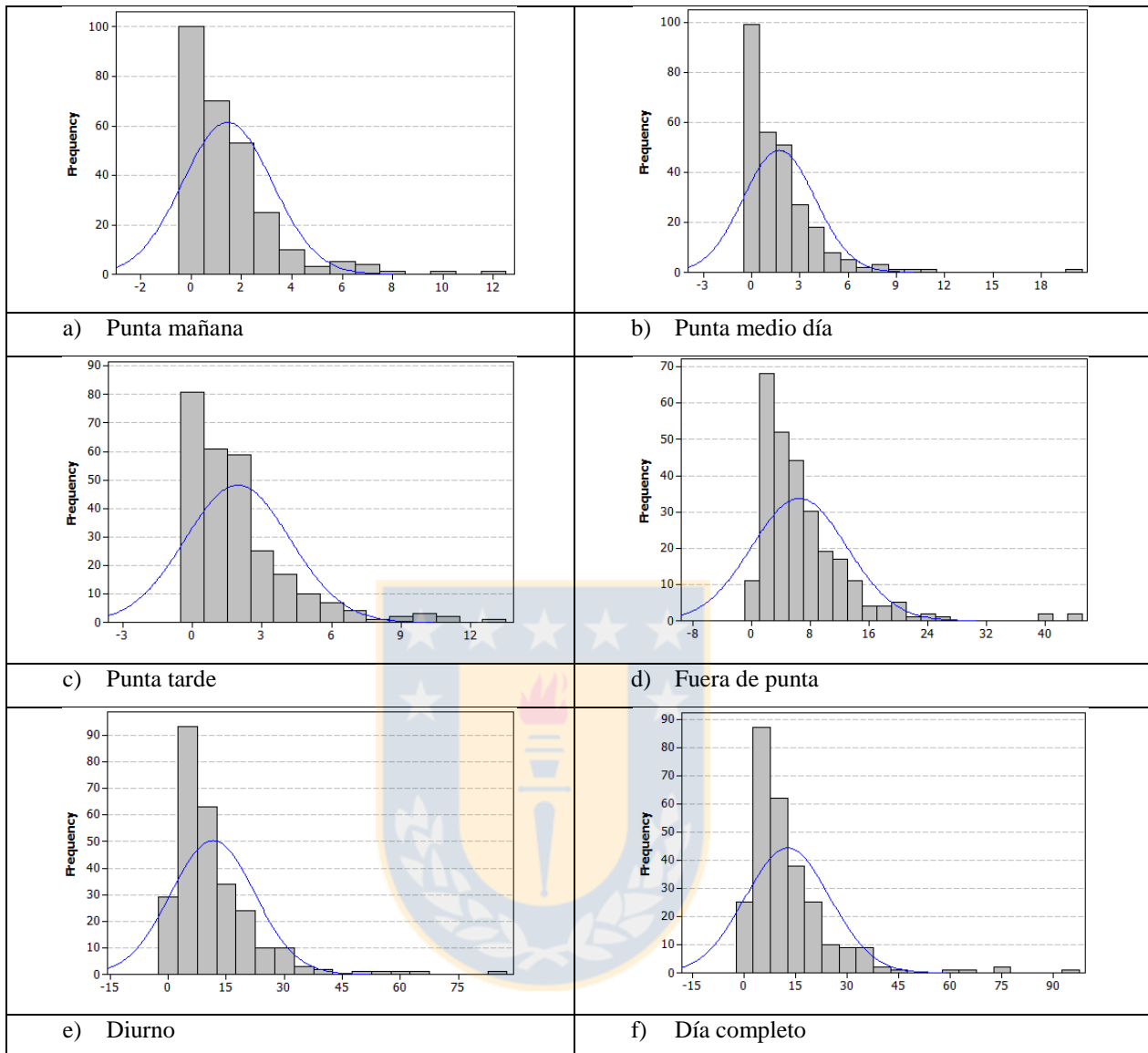


Figura 5.2 Histograma de frecuencia de los accidentes según periodo

En la Figura 5.2 se ajustó por defecto una distribución normal, con la media de accidentes en cada periodo y cómo se puede observar, esta distribución no representa el comportamiento de los datos de accidentes (normalmente las distribuciones que mejor se ajustan a estos datos son la binomial negativa y Poisson). A pesar de esto los test de Anderson-Darling realizados, cumplen con la normalidad a un 95% de confianza. Sin embargo, gráficamente esto no tiene sentido y se observa un sesgo producido por la gran cantidad de ceros en los periodos punta mañana, punta medio día y punta tarde, así como colas más largas con valores extremos. En la Tabla 5.1 se presenta el resultado de la estadística descriptiva del número de accidentes en cada periodo.

Tabla 5.1 Estadística descriptiva de la variable dependiente en cada base de datos

Base de datos	a) PM	b) PMD	c) PT	d) FP	e) Diurno	f) Completo
Cantidad de horas consideradas	2	2	2	10	16	24
Suma total de accidentes del periodo	394	469	533	1761	3157	3431
Media de accidentes del periodo	1.44	1.72	1.95	6.45	11.48	12.57
Media de accidentes por hora	0.72	0.86	0.98	0.65	0.72	0.52
Mediana de accidentes	1	1	1	5	8	9
Número mínimo de accidentes por intersección	0	0	0	0	0	0
Número máximo de accidentes por intersección	12	20	13	44	85	95
Varianza de accidentes del periodo	3.15	4.99	5.10	41.85	117.57	151.27

Las medias y las medianas presentan valores muy cercanos a los valores mínimos. Esto implica que existe una proporción elevada de intersecciones con pocos accidentes (valores en torno a uno). Existen valores máximos de un orden de magnitud mayor a las medias, los cuales se revisarán en el análisis posterior si son *outliers*. La media de accidentes por hora indica que el periodo punta tarde tiene la mayor ocurrencia de accidentes y que los periodos punta mañana, medio día y tarde tienen un mayor promedio de accidentes que el periodo fuera de punta.

5.4 Estadística descriptiva de las variables independientes

La Tabla 5.2 y las Figuras 5.3, 5.4, 5.5 y 5.6 presentan los resultados de frecuencia de las variables de infraestructura.

Tabla 5.2 Frecuencias de las variables del tipo *dummy* de la base de datos

Variables <i>dummy</i> : Presencia de	Sí (%)	No (%)	Variables <i>dummy</i> : Presencia de	Sí (%)	No (%)
Mediana en Rama Principal	56	44	Pista exclusiva giro a derecha en RP	8	92
Mediana en Rama Secundario	9	91	Pista exclusiva giro a derecha en RS	12	88
Pista exclusiva giro a izquierda en RP	24	76	Valla peatonal	41	59
Pista exclusiva giro a izquierda en RS	12	88			

En la Tabla 5.2 se observa que las variables *dummy* que poseen un desequilibrio en las muestras son: la presencia de mediana en la rama secundaria, de pista exclusiva de giro a izquierda en la rama secundaria y la presencia de pista exclusiva de giro a derecha en la rama principal y la secundaria.

Número de ramas de acceso	Resultado										
<table border="1"> <caption>Data for Figure 5.3</caption> <thead> <tr> <th>Nº RAMAS ACCESO</th> <th>Percent</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>1.4652</td> </tr> <tr> <td>2</td> <td>38.8278</td> </tr> <tr> <td>3</td> <td>46.8864</td> </tr> <tr> <td>4</td> <td>12.8205</td> </tr> </tbody> </table>	Nº RAMAS ACCESO	Percent	1	1.4652	2	38.8278	3	46.8864	4	12.8205	<p>El 47% de las intersecciones presentan 3 ramas de acceso, el 39% tienen 2 y el 13% posee 4. El 1% de intersecciones presenta una sola rama de acceso. La mayor proporción de intersecciones poseen entre 2 y 3 ramas de acceso alcanzando un 86% del total de la muestra.</p>
Nº RAMAS ACCESO	Percent										
1	1.4652										
2	38.8278										
3	46.8864										
4	12.8205										

Figura 5.3 Distribución de frecuencia del número de ramas de acceso en las intersecciones

Número de pistas en la rama principal	Resultado																
<table border="1"> <caption>Data for Figure 5.4</caption> <thead> <tr> <th>Nº PISTAS RP</th> <th>Percent</th> </tr> </thead> <tbody> <tr> <td>2</td> <td>21.6117</td> </tr> <tr> <td>3</td> <td>15.7509</td> </tr> <tr> <td>4</td> <td>15.0183</td> </tr> <tr> <td>5</td> <td>9.89011</td> </tr> <tr> <td>6</td> <td>25.641</td> </tr> <tr> <td>7</td> <td>11.7216</td> </tr> <tr> <td>8</td> <td>0.3663</td> </tr> </tbody> </table>	Nº PISTAS RP	Percent	2	21.6117	3	15.7509	4	15.0183	5	9.89011	6	25.641	7	11.7216	8	0.3663	<p>Un 25% de la muestra presenta seis pistas, siendo la configuración de mayor frecuencia. Un 21% de las intersecciones tiene 2 pistas en la rama principal y luego un 15% tienen 3 pistas y otro 15%, 4 pistas.</p>
Nº PISTAS RP	Percent																
2	21.6117																
3	15.7509																
4	15.0183																
5	9.89011																
6	25.641																
7	11.7216																
8	0.3663																

Figura 5.4 Distribución de frecuencia del número de pistas en la rama principal

Número de pistas en la rama secundaria	Resultado																
<table border="1"> <caption>Data for Figure 5.5</caption> <thead> <tr> <th>Nº PISTAS RS</th> <th>Percent</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>4.0293</td> </tr> <tr> <td>1</td> <td>4.3956</td> </tr> <tr> <td>2</td> <td>67.033</td> </tr> <tr> <td>3</td> <td>17.2161</td> </tr> <tr> <td>4</td> <td>5.86081</td> </tr> <tr> <td>5</td> <td>1.0989</td> </tr> <tr> <td>6</td> <td>0.3663</td> </tr> </tbody> </table>	Nº PISTAS RS	Percent	0	4.0293	1	4.3956	2	67.033	3	17.2161	4	5.86081	5	1.0989	6	0.3663	<p>Un 67% de las intersecciones presentan dos pistas en sus calles secundarias y luego, muy bajo este valor, un 17% tienen 3 pistas. Se observa una distribución casi normal con un gran pico.</p>
Nº PISTAS RS	Percent																
0	4.0293																
1	4.3956																
2	67.033																
3	17.2161																
4	5.86081																
5	1.0989																
6	0.3663																

Figura 5.5 Distribución de frecuencia del número de pistas en la rama secundaria

Ángulo de la intersección (°)	Resultado																		
<table border="1"> <caption>Data for Figure 5.6</caption> <thead> <tr> <th>ÁNGULO INTERSECCIÓN (°)</th> <th>Percent</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>2.1978</td> </tr> <tr> <td>20</td> <td>0.3663</td> </tr> <tr> <td>40</td> <td>1.0989</td> </tr> <tr> <td>60</td> <td>1.8315</td> </tr> <tr> <td>80</td> <td>2.9304</td> </tr> <tr> <td>100</td> <td>4.0293</td> </tr> <tr> <td>120</td> <td>10.2564</td> </tr> <tr> <td>140</td> <td>75.0918</td> </tr> </tbody> </table>	ÁNGULO INTERSECCIÓN (°)	Percent	0	2.1978	20	0.3663	40	1.0989	60	1.8315	80	2.9304	100	4.0293	120	10.2564	140	75.0918	<p>Se observa que el 75% de las intersecciones de la muestra presentan una configuración del tipo perpendicular y el 25% tienen un ángulo menor a 90°.</p>
ÁNGULO INTERSECCIÓN (°)	Percent																		
0	2.1978																		
20	0.3663																		
40	1.0989																		
60	1.8315																		
80	2.9304																		
100	4.0293																		
120	10.2564																		
140	75.0918																		

Figura 5.6 Distribución de frecuencia del ángulo de la intersección

En las Figuras 5.7, 5.8, 5.9, 5.10, 5.11 y 5.12 se muestran los resultados de frecuencias de las variables de operación. Los resultados de las variables de tráfico, definidas en 4.4.3 de cada uno de los seis periodos de estudio, se presentan en Anexo A.5.1. De los principales resultados se observó en la variable tráfico medio opuesto, un 80% de las intersecciones que no tiene flujo opuesto. Esto en parte se debe a que en las intersecciones semaforizadas, por lo general, tienen una operación tal que no ocurran conflictos entre movimientos. Se observó también que en el tráfico medio de taxi colectivos existe un 65% de intersecciones que no presentan flujo y el tráfico de camiones presenta un elevado pico en la distribución alcanzando un 40% en torno a los 600 (veh/h). Se observa en general una gran variabilidad en el resto de estas variables de flujo.

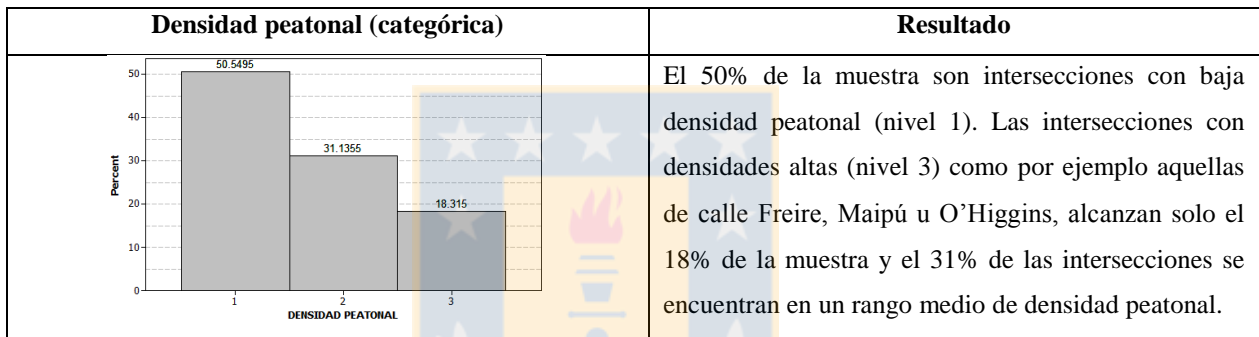


Figura 5.7 Distribución de frecuencia de la densidad peatonal de la intersección

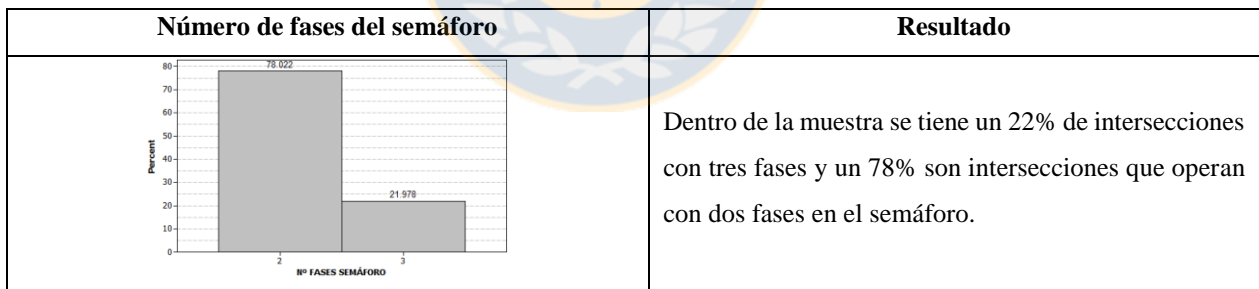


Figura 5.8 Distribución de frecuencia del número de fases del semáforo de la intersección

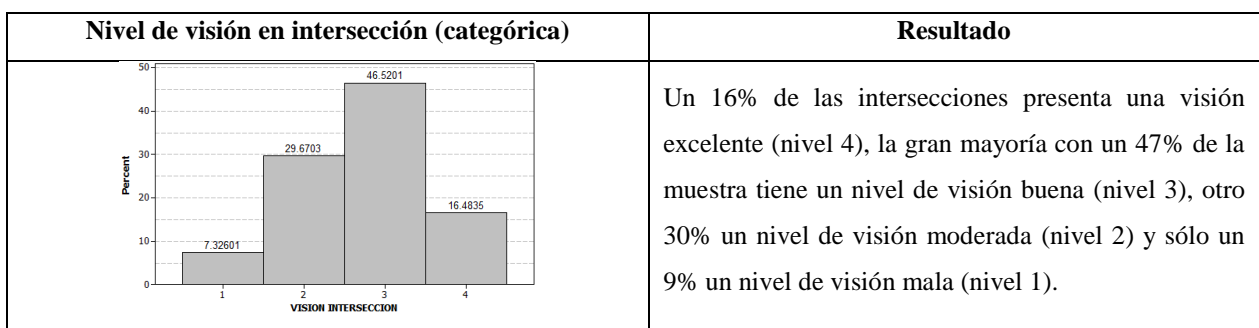


Figura 5.9 Distribución de frecuencia del nivel de visión de la intersección

Visibilidad (m)	Resultado
	<p>Un 80% de las intersecciones poseen entre 0 y 40 m de visibilidad. Los valores de 300 m son para intersecciones peatonales y se asignó por defecto. Un 5% de las intersecciones tiene 200 m de visión el cual se utilizó cuando no existía obstáculo alguno.</p>

Figura 5.10 Distribución de frecuencia de la visibilidad de la intersección

Número de cruces peatonales	Resultado
	<p>El 48% de la muestra tiene 4 cruces peatonales. Un 40% de las intersecciones tienen tres o dos cruces permitidos. Existen casos de intersecciones donde no hay cruces habilitados En cambio, existen valores máximos con ocho cruces peatonales, en del eje Manuel Rodríguez.</p>

Figura 5.11 Distribución de frecuencia de los cruces peatonales de la intersección

Número de cruces peatonales considerando refugio	Resultado
	<p>Se observa una diferencia en la forma de la distribución de la variable con respecto a la anterior. Casi un 80% tienen entre tres y seis cruces peatonales, los valores de esta variable son mayores, pues existen intersecciones, tales como las del eje Paicaví o Los Carrera, en donde las medianas presentes aumentan el número de cruces.</p>

Figura 5.12 Distribución de frecuencia de los cruces peatonales con refugio de la intersección

En las Figuras 5.13, 5.14, 5.15 y 5.16 se muestran los histogramas de frecuencia de las variables relacionadas al entorno.

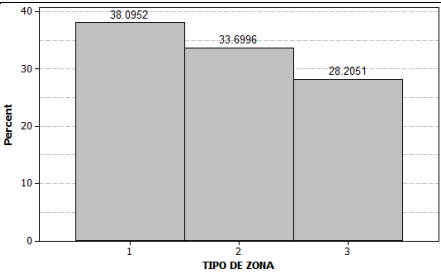
Tipo de zona	Resultado
 <p>A bar chart titled 'Tipo de zona' showing the percentage distribution of three zone types. The y-axis is labeled 'Percent' and ranges from 0 to 40. The x-axis is labeled 'TIPO DE ZONA' with categories 1, 2, and 3. The bars represent: 1 (Residential) at 38.0952%, 2 (Commercial) at 33.6996%, and 3 (Mixed) at 28.2051%.</p>	<p>Esta variable se relaciona con el tipo de uso de suelo predominante en las cercanías de la intersección. Se tiene una distribución de intersecciones, Residencial (1) con un 38%, Comercial (2) con un 34% y Mixto (3) con un 28%.</p>

Figura 5.13 Distribución de frecuencia de tipo de zona de las intersecciones

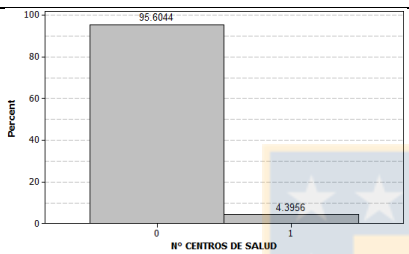
Número de centros de salud	Resultado
 <p>A bar chart titled 'Número de centros de salud' showing the percentage distribution of health centers. The y-axis is labeled 'Percent' and ranges from 0 to 100. The x-axis is labeled 'Nº CENTROS DE SALUD' with categories 0 and 1. The bars represent: 0 centers at 95.6044% and 1 center at 4.3956%.</p>	<p>Existen muy pocas intersecciones con centros de salud cercanos, alcanzando solo un 4% de la muestra.</p>

Figura 5.14 Distribución de frecuencia de número de centros de salud de las intersecciones

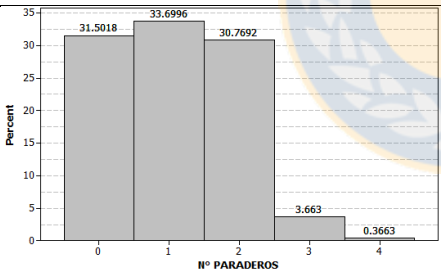
Número de paraderos	Resultado
 <p>A bar chart titled 'Número de paraderos' showing the percentage distribution of bus stops. The y-axis is labeled 'Percent' and ranges from 0 to 35. The x-axis is labeled 'Nº PARADEROS' with categories 0, 1, 2, 3, and 4. The bars represent: 0 stops at 31.5018%, 1 stop at 33.6996%, 2 stops at 30.7692%, 3 stops at 3.663%, and 4 stops at 0.3663%.</p>	<p>De las intersecciones estudiadas, un 31% no presentan paradero, un 34% tienen un paradero y un 31% tienen dos paraderos, el resto con más de dos paraderos alcanza un 4% de la muestra</p>

Figura 5.15 Distribución de frecuencia de número de paraderos de las intersecciones.

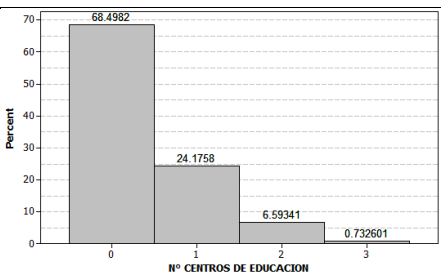
Número de centros de educación	Resultado
 <p>A bar chart titled 'Número de centros de educación' showing the percentage distribution of education centers. The y-axis is labeled 'Percent' and ranges from 0 to 70. The x-axis is labeled 'Nº CENTROS DE EDUCACION' with categories 0, 1, 2, and 3. The bars represent: 0 centers at 68.4982%, 1 center at 24.1758%, 2 centers at 6.59341%, and 3 centers at 0.732601%.</p>	<p>Existen intersecciones sin la influencia de centros de educación, alcanzando un 68%. No obstante, un 24% tiene al menos un establecimiento cercano y un 7% tiene dos.</p>

Figura 5.16 Distribución de frecuencia de número de centros de educación

Se observa que existen variables que concentran gran cantidad de datos en un solo valor, como lo son el número de pistas en la rama secundaria, el ángulo de la intersección, la visibilidad, el número de cruces peatonales o el número de centros de salud. No obstante en otras variables, como el tipo de zona, el número de paraderos o el número de pistas en la rama principal, no existe una clara tendencia en sus distribuciones. De las variables definidas existen algunas que no poseen variabilidad, lo cual se condice con lo graficado. Por lo tanto, dichas variables no podrán explicar adecuadamente el fenómeno en un modelo matemático que represente la ocurrencia de accidentes.

5.5 Análisis de valores extremos

Se presenta el análisis de valores extremos realizado en cada periodo, cuyo objetivo fue identificar aquellas intersecciones que presentan comportamientos extremos con respecto al resto de la muestra, de tal manera de eliminarlas buscando mejorar los resultados de análisis posteriores. Para ello se usó MINITAB 14 (Minitab Inc., 2005) y se obtuvieron los resultados mostrados en la Figura 5.17. Se revisaron los puntos marcados como valores extremos por el *software*, de tal manera de eliminar solo aquellos que realmente se encontraban fuera del orden de magnitud de accidentes de cada periodo.

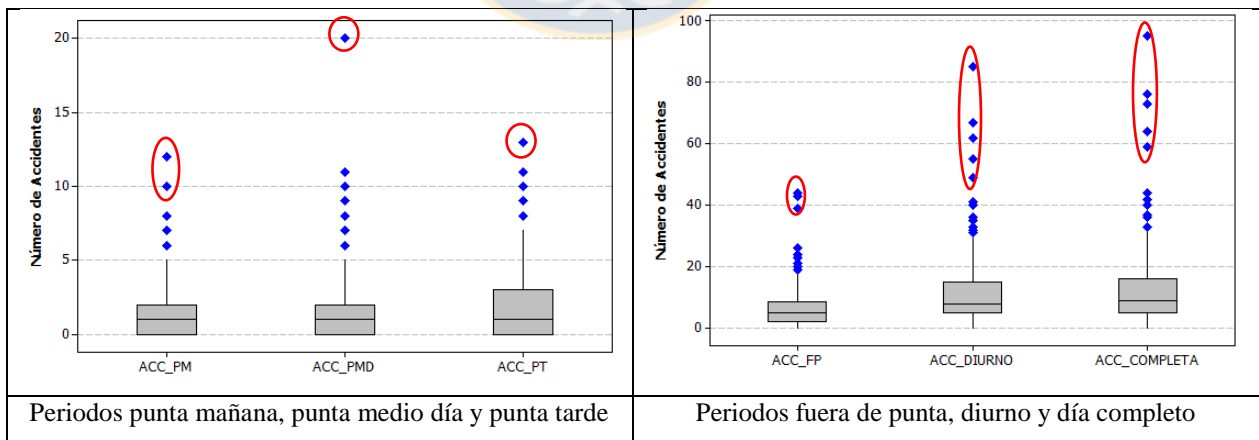


Figura 5.17 Diagramas de caja de accidentes en cada periodo

Como se observa en la Figura 5.17, existen valores extremos en cada base de datos. En este caso los valores representan el número de accidentes, por lo cual deben revisarse las intersecciones que

poseen estos elevados números de accidentes. En la Tabla 5.3 se presentan las intersecciones individualizadas con su número de accidentes, las que se eliminaron en cada base de datos que correspondiese.

Tabla 5.3 Número de accidentes de intersecciones identificadas como valores extremos

Intersecciones	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Av. 21 de mayo - Pedro de Oña	12	20	-	43	85	95
Av. P. A. Cerda - D. Belmar	10	-	-	39	67	73
Av. Los Carrera - Av. Paicaví	-	-	-	44	62	76
Av. Los Carrera - Av. Prat	-	-	13	-	49	59
Av. P. A. Cerda - Diagonal Biobio	-	-	-	-	55	64
Media	1.44	1.72	1.95	6.45	11.48	12.57

La Tabla 5.4 muestra los resultados en los distintos periodos del número de intersecciones, en donde en el periodo diurno y día completo se eliminaron la mayor cantidad de intersecciones, alcanzando cinco unidades, y en el periodo punta tarde y punta medio día solo se eliminó una intersección. Las diferencias en la cantidad de accidentes originales versus la nueva cantidad de accidentes, debido a las intersecciones eliminadas, varió entre un 2,4% en el periodo punta tarde, hasta un 10% cuando se consideró el periodo de día completo.

Tabla 5.4 Composición inicial y post-análisis de valores extremos de las bases de datos

Cantidad	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Intersecciones iniciales	273	273	273	273	273	273
Accidentes iniciales	394	469	533	1761	3157	3431
Intersecciones finales	271	272	272	270	268	268
Accidentes finales	372	449	520	1635	2839	3064
Diferencia de accidentes (%)	5.6	4.3	2.4	7.1	4.7	10.0

5.6 Ajuste de distribuciones de la variable dependiente

El ajuste de distribuciones se realizó en EasyFit 5.5 (MathWave Technologies, 2010). Como se observa en la Tabla 5.5, las distribuciones a las que más se ajustan los datos de accidentes son a la Binomial negativa y a la Poisson. A medida que se aumenta el rango horario de la base de datos, y por consiguiente, la cantidad de datos de accidentes, la distribución Poisson deja de ser adecuada y la Binomial Negativa obtiene un mejor ajuste, el cual se ve reflejado en los valores del estadístico

de Kolmogorov-Smirnov y de Anderson-Darling, en donde mientras menor sean sus valores mayor es el ajuste. Esto se contrasta con los histogramas de la sección 5.3 en donde se intentó ajustar una distribución normal a los datos de accidentes.

Tabla 5.5 Ajustes de distribuciones para la variable de accidentes

Variable	Distribución	Parámetros de la distribución	Estadístico de Kolmogorov Smirnov	Estadístico de Anderson Darling
Accidentes Punta Mañana	1. Poisson	$\lambda=1.4432$	0.23	23.50
	2. Binomial Neg.	$n=3$ $p=0.6751$	0.30	26.19
Accidentes Punta Medio Día	1. Poisson	$\lambda=1.7179$	0.18	20.79
	2. Binomial Neg.	$n=2$ $p=0.5379$	0.28	20.84
Accidentes Punta Tarde	1. Poisson	$\lambda=1.9524$	0.16	18.10
	2. Binomial Neg.	$n=2$ $p=0.5060$	0.25	17.92
Accidentes Fuera de Punta	1. Binomial Neg.	$n=2$ $p=0.2366$	0.10	2.84
	2. Geométrica	$p=0.1342$	0.21	14.12
Accidentes Diurno	1. Binomial Neg.	$n=2$ $p=0.14837$	0.06	1.39
	2. Geométrica	$p=0.08013$	0.17	11.49
Accidentes Día Completo	1. Binomial Neg.	$n=2$ $p=0.1372$	0.08	1.73
	2. Geométrica	$p=0.0737$	0.17	10.39

Según los resultados de la Tabla 5.5 el periodo diurno es en donde se ajusta de mejor manera la distribución binomial negativa, seguido del periodo día completo, pues el estadístico de Kolmogorov-Smirnov es inferior a 0.10. Por otra parte, en los periodos punta tarde, punta medio día y punta mañana la distribución Poisson se ajusta mejor que la binomial negativa según los estadísticos de Kolmogorov-Smirnov y Anderson-Darling. No obstante y según Navarro *et al.* (2001), en datos con sobre-dispersión como es el caso de los accidentes, el análisis con la distribución Poisson conduce a una subestimación de los errores estándares de los coeficientes, pudiendo derivar en la significación estadística de factores que realmente no estén asociados con el fenómeno. En cambio, la binomial negativa si capta parte de la variancia que no identifica la regresión de Poisson. Por consiguiente en estas bases de datos, donde la sobre dispersión y la cantidad de datos con valor 0 aumenta, la distribución binomial negativa disminuye su ajuste y la Poisson aumenta artificialmente.

5.7 Análisis de los efectos principales sobre los accidentes

El análisis de los efectos principales de las variables sobre la media de accidentes busca identificar una tendencia del efecto sobre la media de los accidentes de manera gráfica. Este análisis se realizó

agrupando las variables según sus relaciones causales de tal modo de obtener el efecto relativo de cada una sobre la media de los accidentes. Con esto fue posible determinar las variables que tenían mayor relevancia dentro de cada grupo. Los grupos que se definieron se muestran en la Tabla 5.6.

Por otra parte, para representar en los gráficos el efecto tendencial de las variables continuas sobre la media de accidentes fue necesario categorizarlas con el algoritmo k-medias (Hartigan y Wong, 1979). Gracias a la categorización realizada se pudo obtener una visualización gráfica de la tendencia del efecto de las variables sobre la media. Los resultados gráficos de los efectos principales sobre la media de accidentes de cada grupo de variables y de cada periodo se presentan en el Anexo A.5.2. Además, se realizó un ranking de importancia del efecto sobre la media de cada variable en los grupos, resumido en una tabla que se muestra en Anexo A.5.3. Finalmente, se presenta también en Anexo A.5.3 una tabla que resume individualmente el comportamiento de la tendencia de cada efecto sobre la media de accidentes en cada periodo.

Tabla 5.6 Grupos de variables para el análisis de efectos principales sobre la media

Variables del flujo por movimientos	Variables del flujo por tipología	Variables Peatonales	Variables relacionadas a los giros	Variables del Diseño Geométrico	Variables de la Visibilidad	Otras Variables
- TMD Intersección	- TMD Vehículos Livianos	- Valla Peatonal	- Pista Exclusiva Giro Izquierda Rama Principal	- N° de Accesos en intersección	- Ángulo de la intersección	- Tipo intersección
- TMD Rama Principal	- TMD Camiones	- Mediana RP	- Pista Exclusiva Giro Izquierda Rama Secundaria	- N° de Pistas Rama Principal		
- TMD Rama Secundaria	- TMD Taxi colectivos	- Mediana RS	- Pista Exclusiva Giro Derecha Rama Principal	- N° de Pistas Rama Secundaria	- Nivel de visión Intersección	- N° fases semáforo
- TMD Giro Derecha	- TMD Taxi buses	- Densidad Peatonal	- Pista Exclusiva Giro Derecha Rama Secundaria			
- TMD Giro Izquierda	- TMD No motorizado	- N° Cruces peatonales	- Giro Izquierda RP	- N° calles unidireccionales	- Visibilidad	- Tipo de zona
- TMD Flujo Opuesto	- TMD Buses	- N° Cruces peatonales con refugio	- Giro Izquierda RS			
		- N° Paraderos				
		- N° Educación				
		- N° Salud				

Nota: TMD de aquí en adelante es el tráfico medio diario.

Del análisis se concluyó que de las variables relacionadas con el peatón, la que tiene un mayor efecto en la ocurrencia de accidentes es el número de paraderos y en segundo lugar el número de cruces peatonales con refugio; esto ocurrió en todos los periodos. En las bases PT, PMD, FP, Diurno y día completo la variable número de centros de educación genera un aumento de accidentes

considerable a medida que ésta aumenta, no así en el periodo PM, en donde no genera ningún efecto. En el grupo de variables relacionadas a los giros, la variable pista exclusiva de giro a izquierda en la rama principal resultó ser la de mayor efecto sobre la media para todas las bases de datos excepto en el periodo PM.

En lo que respecta al diseño geométrico, la variable de mayor importancia en todos los periodos es el número de pistas en la rama secundaria. Aquella que no genera ningún efecto sobre la media de accidentes es el número de calles unidireccionales. En el grupo de variables que guardan relación con la visibilidad, la variable que mide este atributo en metros resultó tener mayor efecto sobre la media de accidentes en todos los periodos.

Finalmente, en el grupo Otras Variables, el tipo de zona y número de fases del semáforo poseen un efecto sobre la media de accidentes. En todos los periodos revisados, el número de fases del semáforo genera un mayor aumento en la media de accidentes. Con respecto al tipo de zona, ésta genera un efecto de menor magnitud en la media de accidentes en todos los periodos a excepción del PM.

5.8 Análisis bi-variado

El análisis bi-variado permitió revisar cuantitativamente los efectos de cada variable independiente sobre el número de accidentes ocurridos en cada intersección. Para las variables continuas se usó la correlación de Pearson, para las variables de dos niveles se utilizó la prueba t-Student, mientras que para las variables de más de dos niveles se usó el ANOVA.

5.8.1 Análisis de variables continuas

El análisis bi variado se realizó para todas las variables continuas de la base de datos. Los resultados de significancia estadística son considerando un intervalo de confianza del 95% ($p \leq 0.05$). En las

Tablas 5.7 y 5.8 se muestran los resultados de la correlación de Pearson calculada de cada una de estas variables con respecto al número de accidentes.

Tabla 5.7 Correlación de Pearson del número de accidentes con las variables continuas en Día completo, Diurno y FP

Variables independientes continuas	Periodo Día Completo		Periodo Diurno		Periodo Fuera de Punta	
	Valor p	Correlación de Pearson	Valor p	Correlación de Pearson	Valor p	Correlación de Pearson
Visibilidad	-		-		-	
TMDRS/TMDINTER	-		-		-	
TMDRS/TMDRP	-		-		-	
TMD_NOMOTOR	0.000	0.393	0.000	0.381	0.000	0.354
TMD_BUS	0.000	0.366	0.000	0.352	0.000	0.315
TMD_TXB	0.000	0.387	0.000	0.384	0.000	0.340
TMD_TXC	-		-		-	
TMD_CAM	0.000	0.476	0.000	0.464	0.000	0.476
TMD_VL	0.000	0.552	0.000	0.532	0.000	0.523
TMD_GI	0.000	0.83	0.000	0.270	0.000	0.235
TMD_GD	0.000	0.364	0.000	0.354	0.000	0.342
TMD_OP	-		-		-	
TMD_RS	0.000	0.341	0.000	0.329	0.000	0.298
TMD_RP	0.000	0.506	0.000	0.493	0.000	0.488
TMD_INTER	0.000	0.591	0.000	0.575	0.000	0.559

Tabla 5.8 Correlación de Pearson del número de accidentes con las variables continuas en PM, PMD y PT

Variables independientes continuas	Periodo Punta Mañana		Periodo Punta Medio Día		Periodo Punta Tarde	
	Valor p	Correlación de Pearson	Valor p	Correlación de Pearson	Valor p	Correlación de Pearson
Visibilidad	-		-		-	
TMDRS/TMDINTER	-		-		-	
TMDRS/TMDRP	-		-		-	
TMD_NOMOTOR	0.000	0.239	0.000	0.250	0.000	0.292
TMD_BUS	0.000	0.215	0.000	0.293	0.000	0.292
TMD_TXB	0.000	0.334	0.000	0.297	0.000	0.292
TMD_TXC	-		-		-	
TMD_CAM	0.000	0.399	0.000	0.268	0.000	0.248
TMD_VL	0.000	0.375	0.000	0.361	0.000	0.417
TMD_GI	0.001	0.194	0.000	0.228	0.000	0.256
TMD_GD	0.000	0.268	0.000	0.260	0.000	0.250
TMD_OP	-		0.049	0.119	-	
TMD_RS	0.000	0.223	0.000	0.260	0.000	0.307
TMD_RP	0.000	0.368	0.000	0.330	0.000	0.354
TMD_INTER	0.000	0.419	0.000	0.440	0.000	0.432

Nota: cuando no existe valor p es porque no resultó significativo estadísticamente.

Se observa que la variable visibilidad no resulta significativa en ningún periodo. Esto indica que la manera de obtener el valor de visibilidad puede no estar considerando lo que en realidad tienen como visión los conductores de las intersecciones. Otras variables como el tráfico medio de taxi

colectivos (TXC) así como las relaciones entre los flujos de la rama secundaria con la intersección y rama principal no resultan significativos y esto se debe a la existencia de pocos datos; por ejemplo, los TXC están presentes en pocas intersecciones. La variable del tráfico medio opuesto resulta solo significativa estadísticamente en el periodo PMD ($P=0.049 < 0.05$). Luego, esta variable debe tratarse con cuidado y buscar una manera de mejorar su ajuste si se desea utilizar, pues con los resultados no se puede asegurar una relación causal.

En resumen, se observan relaciones con coeficientes positivos cuyos valores van desde 0.12 hasta 0.59 todos con $p < 0.05$. Otro aspecto que debe notarse es el peso relativo de las variables al comparar entre cada base de datos. Se observa que, a medida que el periodo de horas de accidentes aumenta, variables más generales, como el flujo total de la intersección (TMD_INTER), tienen un mayor peso, con correlaciones de Pearson de 0.56 a 0.59 en los periodos día completo, diurno y FP, mientras en las bases de datos de dos horas (PM, PT y PMD) el valor de este peso disminuye a ordenes de 0.42 a 0.44.

Este efecto ocurre, análogamente, en las variables tráfico total de vehículos livianos (TMD_VL) y tráfico total de camiones (TMD_CAM), por lo cual a medida que se aumenta el periodo representativo, resulta mejor usar este tipo de variables más agregadas en un futuro modelo.

En conclusión, la visibilidad (m) no resultó ser significativa estadísticamente en ningún caso. No obstante, las variables de tráfico, a excepción del flujo de taxi colectivos y del flujo opuesto, resultaron tener una correlación estadística en la ocurrencia de accidentes en todos los periodos.

5.8.2 Análisis de variables con dos niveles

La prueba t-student estudia la relación entre una variable dependiente cuantitativa y una variable con dos grupos o niveles. En la Tabla 5.9 se presentan las medias de accidentes en cada nivel de las variables analizadas.

La variable número de fases del semáforo presenta en las intersecciones con tres fases mayores niveles de accidentes que aquellas con dos fases, en todos los periodos. Del mismo modo, cuando existen giros a la izquierda tanto en la rama principal como en la rama secundaria, la media de accidentes es mayor. Para las variables que tienen relación con las pistas exclusivas de giro, las medias de accidentes son mayores cuando hay presencia de estas pistas. Esto se explica, porque al ser intersecciones más complejas, donde además el flujo que gira es elevado (como consecuencia existe la pista exclusiva de viraje), existen mayores conflictos entre movimientos en la operación de la intersección. Con esto se puede ver que medidas como las pistas de viraje exclusivas no son suficientes para disminuir el número de accidentes en intersecciones complejas. Finalmente, en variables como el número de centros de salud, se observa que no existe ninguna tendencia.

Tabla 5.9 Media de accidentes de cada nivel de cada variable dicotómica

Variables	Niveles	Número	Media de Accidentes					
			PM	PMD	PT	FP	DIURNO	DÍA COMP.
Pista exc giro der rama principal	Sí	23	2.43	2.35	2.22	6.20	15.91	17.57
	No	250	1.35	1.66	1.93	9.17	11.07	12.11
Pista exc giro der rama secundaria	Sí	33	2.36	2.06	1.79	8.52	14.45	16.55
	No	240	1.32	1.67	1.98	6.17	11.07	12.02
Pista exc giro izq rama principal	Sí	66	2.06	2.36	2.52	8.70	15.47	17.44
	No	207	1.25	1.51	1.77	5.73	10.21	11.01
Pista exc giro izq rama secundaria	Sí	34	2.18	1.82	2.09	7.59	13.50	15.12
	No	239	1.34	1.70	1.93	6.29	11.19	12.21
Mediana rama principal	Sí	152	1.72	1.91	2.33	7.57	13.39	14.85
	No	121	1.10	1.47	1.48	5.05	9.07	9.70
Mediana rama secundaria	Sí	25	2.08	1.88	2.04	8.20	14.04	15.96
	No	248	1.38	1.70	1.94	6.27	11.22	12.23
Valla peatonal	Sí	112	1.61	1.95	2.17	7.16	12.79	14.09
	No	161	1.33	1.56	1.80	5.96	10.57	11.51
Giro izq rama principal	Sí	132	1.51	1.84	1.99	6.70	11.93	13.29
	No	141	1.38	1.60	1.91	6.22	11.06	11.89
Giro izq rama secundaria	Sí	192	1.56	1.83	1.97	6.87	12.14	13.33
	No	81	1.16	1.44	1.91	5.46	9.91	10.77
Número fases semáforo	2	213	1.20	1.50	1.77	5.60	10.00	10.77
	3	60	2.32	2.48	2.62	9.47	16.72	18.95
Centro de salud	0	261	1.44	1.71	1.96	6.52	11.54	12.66
	1	12	1.42	1.92	1.83	5.00	10.08	10.58

En la Tabla 5.10 se presentan los resultados de los valores p de los test t de cada variable analizada, solo cuando resultaron significativas y cumplían con la prueba de Levene.

Tabla 5.10 Significancia estadística según la prueba t-Student para cada variable de dos niveles

Variables	Pruebas	Significancias					
		ACC PM	ACC PMD	ACC PT	ACC FP	ACC DIURNO	ACC COMP
Pista exc giro der rama principal	Levene	0.000	0.339	0.030	0.000	0.000	0.000
	T	-	-	-	-	-	-
Pista exc giro der rama secundaria	Levene	0.002	0.489	0.448	0.004	0.028	0.007
	T	0.015	-	-	-	-	-
Pista exc giro izq rama principal	Levene	0.000	0.003	0.007	0.000	0.000	0.000
	T	0.012	0.044	0.040	0.013	0.009	0.005
Pista exc giro izq rama secundaria	Levene	0.052	0.951	0.061	0.148	0.082	0.092
	T	0.030	-	-	-	-	-
Mediana rama principal	Levene	0.007	0.733	0.008	0.000	0.000	0.000
	T	0.003	-	0.001	0.001	0.001	0.000
Mediana rama secundaria	Levene	0.030	0.782	0.072	0.102	0.082	0.049
	T	-	-	-	-	-	-
Valla peatonal	Levene	0.006	0.013	0.323	0.081	0.026	0.016
	T	-	-	-	-	-	-
Giro izq rama principal	Levene	0.065	0.208	0.180	0.043	0.025	0.018
	T	-	-	-	-	-	-
Giro izq rama secundaria	Levene	0.169	0.262	0.630	0.468	0.276	0.229
	T	0.068	-	-	-	-	-
Número fases semáforo	Levene	0.000	0.000	0.000	0.000	0.000	0.000
	T	0.001	0.003	0.010	0.003	0.003	0.001
Centro de salud	Levene	0.983	0.950	0.807	0.107	0.522	0.453
	T	-	-	-	-	-	-

De la Tabla 5.10 se tiene como resultado que la variable que mejor se comporta es la del número de fases del semáforo, en donde tanto el comportamiento de la media como la significancia estadística en las seis bases de datos son consistentes. Del mismo modo, la variable asociada a la presencia de pista exclusiva de giro a izquierda en la rama principal resulta significativa para las seis bases de datos también. La variable de presencia de mediana en rama principal resulta significativa en cinco BD, a excepción de aquella del periodo PMD. Por otra parte, existen algunas variables que no se comportan adecuadamente, siendo solo significativas para la BD PM, como son la pista exclusiva giro derecha/izquierda en rama secundaria y el giro a izquierda en rama secundaria.

5.8.3 Análisis de las variables de tres o más niveles

Para aquellas variables con más de dos niveles se realizó el análisis ANOVA. En éste se comparan las varianzas de cada variable de tal modo de encontrar si son o no estadísticamente diferentes. Para ello, se debe cumplir la homogeneidad de varianzas. Cuando esto se cumple se utiliza la tabla

ANOVA. Cuando no es el caso, se realizan las pruebas robustas de igualdad de medias de Welch y posteriormente se realizan las pruebas post hoc de Bonferroni o Games Howel.

Tabla 5.11 Resultados del ANOVA de las variables de más de dos niveles.

Pruebas Variables	Homogeneidad de varianzas						Significancia estadística					
	PM	PMD	PT	FP	DIURNO	COMP	PM	PMD	PT	FP	DIURNO	COMP
TMD Intersección	No	No	No	No	No	No	0.000	0.000	0.000	0.000	0.000	0.000
TMD Rama principal	No	No	No	No	No	No	0.000	0.000	0.000	0.000	0.000	0.000
TMD Rama secundaria	Si	Sí	No	No	Sí	No	0.001	0.000	-	0.039	0.003	0.017
TMD Opuesto	No	No	Sí	No	No	No	-	0.000	-	-	-	-
TMD Giro derecha	No	No	No	No	No	No	-	0.000	0.048	0.001	0.003	0.002
TMD Giro izquierda	No	No	No	No	No	No	-	-	-	0.014	0.024	0.017
TMD Vehículos Liv.	No	No	No	No	No	No	0.000	0.000	0.000	0.000	0.000	0.000
TMD Camiones	No	No	No	No	No	No	0.000	0.005	0.001	0.000	0.000	0.000
TMD Taxi colectivos	Sí	No	No	Sí	Sí	Sí	-	-	-	-	-	-
TMD Taxi buses	No	No	No	No	No	No	0.000	0.000	0.000	0.000	0.000	0.000
TMD Buses	No	No	No	No	No	No	0.003	0.038	0.006	0.002	0.000	0.000
TMD No motorizado	No	Sí	No	No	No	No	0.021	0.000	0.001	0.000	0.000	0.000
TMDRS / TMDRP	Sí	Sí		No	No	No	-	0.021		0.008	0.002	0.001
TMDRS / TMDINTER	Sí	Sí		Sí	No	No	-	0.012		0.014	0.011	0.009
N° Cruces peatonales	No	Sí	Sí	No	No	No	-	-	-	0.002	-	-
N° Cruces peatonales con refugio peatonal	No	Sí	Si	No	No	No	0.005	-	-	0.004	0.008	0.008
Giro izq. RP Pista Exc.	No	No	No	No	No	No	0.018	-	-	0.027	0.014	0.011
Giro izq. RS Pista Exc.	No	Sí	Sí	Sí	Sí	Sí	0.049	-	-	-	-	-
Densidad Peatonal	Sí	No	Sí	Sí	Sí	Sí	-	0.048	-	0.049	0.051	-
Visibilidad	No	Sí	No	No	No	No	0.006	-	-	0.000	0.001	0.001
Nivel de visión	Sí	Sí	Sí	Sí	Sí	Sí	-	-	-	-	-	-
N° calles unidireccional	No	No	Sí	No	No	No	0.001	0.002	-	0.035	0.007	0.001
Angulo obtuso	No	No	Sí	No	No	No	0.000	0.011	-	0.000	0.000	0.000
N° Pista Rama principal	No	No	No	No	No	No	0.000	0.001	0.000	0.000	0.000	0.000
N° Pista Rama secundaria	No	No	No	No	No	No	0.000	0.002	0.015	0.000	0.000	0.000
N° Paraderos	No	No	No	No	No	No	0.000	0.001	0.002	0.000	0.000	0.000
N° Centros de educación	Sí	Sí	Sí	No	Sí	Sí	-	-	-	0.024	-	-
Tipo de Zona	No	No	Sí	Sí	Sí	No	-	-	-	-	-	-

La Tabla 5.11 muestra los resultados del análisis ANOVA, para las variables relacionadas con la operación, infraestructura y entorno. Los resultados completos de cada análisis realizado en SPSS Statistics 22.0 (IBM, 2014) se encuentran en el Anexo Magnético Capítulo 5.

Es importante notar que para este experimento, donde se cuenta con 273 intersecciones, la categorización de variables (número de pistas por rama, número de paradero, número de accesos, ángulo, los cruces peatonales, visibilidad y las variables de flujos) resulta significativa para buscar mejorar los comportamientos y resultados experimentales, pues se disminuye la cantidad de niveles analizados.

Se observa que las variables de flujo vehicular tienen una significancia estadística importante en su relación con el número de accidentes en todos los periodos (valor $p < 0.05$), no así para variables como el tipo de zona, el número de centros de educación o la presencia de giro a izquierda con pista exclusiva.

En el análisis ANOVA se categorizaron las variables continuas (Flujos y visibilidad), las cuales fueron anteriormente analizadas según la correlación de Pearson, en la sección 5.8.1. Comparando los resultados de ambos análisis, solo se encontró una diferencia importante con respecto a la variable visibilidad en donde según el ANOVA resulta significativa. Con estos resultados es posible decir con mayor certeza que las variables relacionadas al flujo son muy importantes para la elaboración de un modelo de accidentes, pues resultaron significativas en casi todas sus formulaciones, en todos los periodos analizados y según los distintos análisis realizados.

5.8.4 Síntesis de los resultados

Finalmente se muestran las variables que sí poseen un efecto en la ocurrencia de accidentes de todas aquellas definidas inicialmente. Se presenta en la Tabla 5.12 el listado de variables que resultaron significativas y consistentes en los distintos análisis que se realizaron.

Tabla 5.12 Resumen de variables significativas con el n° de accidentes.

Grupo de variables	Variable	Grupo de variables	Variable
Operación	TMD Intersección	Infraestructura	N° de calles unidireccionales
	TMD Rama principal		Ángulo de intersección
	TMD Rama secundaria		N° de pistas de rama principal
	TMD Giro derecha		N° de pista de rama secundaria
	TMD Vehículos livianos		Pista exclusiva giro izq. en rama principal
	TMD Camiones		N° de pista de rama secundaria
	TMD Taxi buses		Pista exclusiva giro izq. en rama principal
	TMD Buses		N° de pista de rama secundaria
	TMD No motorizado		Pista exclusiva giro izq. en rama principal
	TMDRS/TMDRP		Mediana en rama principal
	TMDRS/TMD Intersección		
	N° Cruces peatonales con refugio		
	Giro izquierda en rama principal con pista exclusiva		
	Densidad peatonal		
	Visibilidad		
	N° de fases de semáforo	Entorno	N° de Paraderos

5.9 Conclusiones

Del análisis exploratorio se obtuvo el comportamiento en la base de datos de las variables independientes, en donde se pudo ver cuáles tenían mayor variabilidad. Del mismo modo, se observó que, aunque los resultados del test de normalidad de Anderson-Darling fueran positivos, gráficamente los accidentes no tienen una distribución normal. En efecto, según lo obtenido del ajuste de distribuciones a los datos de accidentes, la distribución binomial negativa y la Poisson son las que mejor se ajustan a los accidentes en los periodos analizados. Existen en cada periodo ciertas intersecciones con valores elevados del número de accidentes en comparación a la media. Estos datos se analizaron de tal modo de sacarlos de la muestra para evitar sesgos o distorsiones en los resultados.

Del listado de variables que se dispuso, un número importante no resultaron significativas a la hora de compararlas y buscar relaciones con respecto a los accidentes en las intersecciones. En las variables relacionadas a la operación, de 19 iniciales resultaron ser 16 significativas

estadísticamente (84%). En aquellas relacionadas con la infraestructura, de 13 variables definidas solo seis variables resultaron significativas estadísticamente (46%) y, finalmente, de las cuatro variables definidas relacionadas al entorno, solo una resultó significativa (25%).

Cuando las variables analizadas tienen poca variabilidad o en muchas intersecciones los valores son cero, existe una tendencia a que no resulten significativas como fue el caso de la variable tráfico medio de taxi colectivos o la variable tráfico medio opuesto. Esto implica que en dichas variables no se puede ver su impacto sobre el número de accidentes y, por consiguiente, no se pueden agregar a un modelo.

El análisis estadístico realizado (correlación de Pearson, test t-Student, ANOVA) permite tener una idea base del efecto y de las relaciones de las variables con el número de accidentes. De este modo, es posible saber cuáles incorporar a un futuro modelo de estimación de accidentes en intersecciones urbanas.



CAPÍTULO 6 MODELACIÓN DE ACCIDENTES

En este capítulo se presentan los resultados relativos a la modelación que se realizó con la información de base de las variables obtenidas como significativas en el análisis realizado anteriormente y presentado en el capítulo 5.

6.1 Metodología de análisis

A continuación se presenta la metodología de trabajo utilizada para realizar el análisis de los modelos generados y la convergencia hacia un modelo con el mejor ajuste y significación física y estadística. Dicho análisis se realizó con *software* estadístico y se procedió de la siguiente manera:

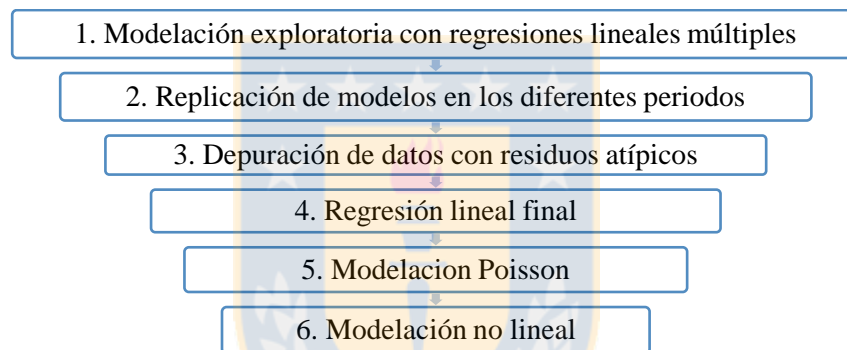


Figura 6.1 Esquema de la metodología de la modelación.

6.1.1 Modelación exploratoria de regresiones lineales múltiples

Un modelo de regresión lineal tiene como base los supuestos de linealidad, normalidad y homocedasticidad. A pesar de estos alcances propios de la regresión lineal, se consideró pertinente realizar la primera aproximación de modelos y de patrones de causalidad de accidentes, lo cual se refuerza con la bibliografía revisada donde se utilizaron regresiones en los análisis (Wang *et al.* 2011, Salifu 2003). Por consiguiente, la forma funcional de los modelos analizados se presenta en la Ecuación 6.1.

$$Y = \theta_0 + \sum_{i=1}^n \theta_i x_i \quad (6.1)$$

donde Y es el número de accidentes total en cada intersección en el periodo de análisis, x_i son las variables independientes incluidas ya sea de operación, infraestructura y/o entorno y finalmente θ_i representa los coeficientes de regresión asociados a cada variable x_i .

En esta etapa se utilizó como información de base los resultados obtenidos en el capítulo 5, es decir, se consideraron sólo aquellas variables que resultaron significativas del análisis tanto de los efectos principales sobre la media como del análisis bi-variado. Para efectos prácticos se trabajó con la base de datos del periodo día (diurno). Para encontrar las combinaciones que mejor representaran el número de accidentes de cada intersección se realizaron 46 pruebas con distintas combinaciones de variables (cinco a ocho variables independientes en cada modelo) y considerando en cada forma conceptual definida una combinación lógica y físicamente adecuada entre variables del entorno, de infraestructura y de operación.

Finalizado el análisis se escogieron siete modelos que se consideraron los más interesantes para el análisis. Cinco de estos modelos correspondieron a aquellos con mejores ajustes de R^2 ajustado. Los otros dos modelos se escogieron, pues representan respectivamente, el mejor modelo estimado considerando solo variables de flujo y el mejor modelo estimado considerando cualquier variable, excluyendo aquellas de flujo.

6.1.2 Determinación de formas conceptuales adecuadas

En esta etapa, habiendo analizado los modelos desde un punto de vista de sus ajustes de R^2 ajustado, así como del análisis de las variables significativas (signos de los coeficientes, test t y valores p) se realizó una replicación de los siete modelos escogidos en cada uno de los periodos de análisis, es decir, en los periodos punta mañana (PM), punta medio día (PMD), punta tarde (PT), fuera de punta (FP) y día completo (COMP). El objetivo principal de esta tarea fue analizar y verificar cuan representativos son los modelos al buscar estimar el número de accidentes en periodos más específicos del día, así como en el día completo.

6.1.3 Depuración de datos con residuos atípicos

Desde esta etapa en adelante los análisis y los resultados se realizaron solo en el periodo diurno. De manera tal de mejorar los ajustes, así como de disminuir la dispersión y distorsiones ocasionadas por ciertas intersecciones (valores extremos) se procedió iterativamente a eliminar aquellas intersecciones que generaran residuos atípicos en los resultados de la modelación. En consecuencia se eliminaron de la base de datos éstas intersecciones resultando así la BD del periodo diurno definitiva.

6.1.4 Regresión lineal final

En esta etapa se replicaron los modelos de regresiones estimados en la etapa 2 (sección 6.1.2) y se escogió el modelo que cumpliera con un mejor ajuste R^2 ajustado, con variables significativas (signos lógicos, test $t > 1.75$ y valor- $p < 0.1$, considerando un 90% de confianza) así como también buscando reducir el número de variables utilizadas, sin perjuicio de la predicción del modelo.

6.1.5 Modelación Poisson

En esta etapa se utilizó como base los resultados de los siete modelos analizados y definidos en la etapa 2 (sección 6.1.2). Con esto se procedió a realizar las regresiones pertinentes con la herramienta de regresiones del tipo Poisson. Con este tipo de modelación se buscó ver como incidía el utilizar este tipo de distribuciones que se ajustan de mejor manera a los datos de accidentes en los resultados de los modelos. Se buscó al igual que en la etapa 4 encontrar un modelo único que cumpliera con ser el de mejor ajuste de R^2 así como obtener resultados consistentes de la significancia estadística de las variables utilizadas apuntando también a disminuir el número de variables sin perjuicio de la capacidad estimativa del modelo.

6.1.6 Modelación no lineal

En esta etapa se utilizó información de base de los resultados obtenidos de las etapas anteriores. Con ello se realizaron algunas pruebas de formas funcionales no lineales tomando como base algunos modelos de la literatura como los de Abdel y Keller (2004), AASHTO (2010), Chen

(2016), entre otros. Adecuándose a las variables propias de la investigación en este trabajo se estimaron modelos no lineales. En este sentido se simplificaron los modelos en cuanto a la cantidad de variables y se encontró que la forma funcional exponencial (modelo de exposición al riesgo) resultó ser la mejor en cuanto a la capacidad de estimar el número de accidentes en las intersecciones.

6.2 Modelación lineal

En esta sección se presentan los análisis y resultados obtenidos de la modelación lineal. Es decir, se exponen aquellas combinaciones de variables que resultaron con los mejores ajustes, así como los resultados en los diferentes periodos y finalmente, la modelación lineal del tipo Poisson.

6.2.1 Modelos de regresión lineal múltiple

Como se mencionó en la sección 6.1.1 se realizaron pruebas con distintas combinaciones de variables explicativas, buscando aquella que generase modelos con ajustes de R^2 adecuados (según lo revisado en la literatura, en torno a 0.5) y con las combinaciones de variables que resultaban estadísticamente significativas. Según lo explicado en la sección 6.1.2, se obtuvieron siete formas conceptuales (selección de variables independientes) con las que se realizó el posterior análisis.

En la Tabla 6.1 se presenta la combinación de variables que fueron aquellas de mayor significancia estadística e importancia en el proceso de modelación, pues al realizar las pruebas siempre resultaron significativas y además fueron aquellas que maximizaron el ajuste de los modelos para explicar los accidentes llegando en torno al 40% de ajuste.

Los resultados de cada modelo planteado en lo que respectan los valores del test t y del valor P de cada variable definida se encuentran en el Anexo 6.1.

Tabla 6.1 Modelos analizados

Variables	Formas conceptuales de los modelos analizados						
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7
Constante	x	x	x	x	x	x	x
N° Pistas RP						x	
N° Pistas RS	x	x	x	x	x	x	
Angulo Cat	x	x	x	x	x		
Densidad Peatonal	x	x	x	x	x	x	
Visibilidad Cat	x	x	x			x	
N° Paraderos					x	x	
N° Calles unidireccionales						x	
N° Cruces Peatonales (2)				x			
TMD Intersección	x		x	x	x		
TMD Rama Secundaria							x
TMD Giro Derecha	x	x		x	x		x
TMD Vehículos Livianos		x					x
TMD Camiones	x	x	x	x	x		x
TMD Taxi Buses	x	x	x	x	x		x
TMD Buses							x
TMD No Motorizado							x
TMD RS / TMD RP							x
R2 ajustado (%)	40.91	40.89	40.55	40.81	40.87	27.56	37.65

De la Tabla 6.1 se desprende que el modelo 1 es el con mejor ajuste estadístico alcanzando casi un 41%. No obstante y al revisar los test t de cada una de sus variables se observa que el de la variable TMD Giro a Derecha y de la Visibilidad resultan menores a 1.645 (90% de confianza). Esto implica que dichas variables debiesen eliminarse del modelo. En este sentido en los modelos 2 y 3 ocurre lo mismo con la variable de Visibilidad y en el modelo 2, ocurre también con la variable de TMD Giro a Derecha, no obstante a ésta última en la forma conceptual 4 resulta significativa estadísticamente a un 90% de confianza.

Otro resultado a destacar es la presencia de la constante en cada modelo de regresión, la cual en todos los modelos resultó estadísticamente significativa y el ajuste resultaba mejor. Esto demuestra

y refuerza la idea que en la explicación del número de accidentes en intersecciones urbanas existen muchos otros factores que influyen además de los captados y medidos con las variables definidas.

En lo que respecta el modelo con variables de infraestructura y operación sin considerar aquellas de flujo, se obtuvo que las variables que mejor explican la ocurrencia de accidentes son el número de pistas tanto en la rama principal como secundaria, la densidad peatonal, la visibilidad categorizada, el número de paraderos y el número de calles unidireccionales. No obstante con este tipo de variables solo es posible explicar un 28% el fenómeno. En contraposición, al realizar el mismo ejercicio considerando solo variables de flujo de la intersección (ver modelo 7) se alcanza un R^2 ajustado de 38%. La combinación de variables que logra una mejor explicación del número de accidentes es utilizando el flujo de la rama secundaria, el flujo que gira a derecha, el flujo de camiones, el flujo de taxibuses, el flujo de buses, el flujo no motorizado y finalmente la razón de flujo entre el de la rama secundaria y primaria. Si esto se compara con el modelo completo (ver modelos 1, 2, 3, 4 y 5), en que se consideran las variables de todo tipo, solo se mejora hasta un 41% el ajuste, lo cual resulta marginal.

6.2.2 Modelación en los diferentes periodos

En esta etapa, los modelos obtenidos del periodo diurno se replicaron en los periodos punta mañana, punta medio día, punta tarde, fuera de punta y día completo. Con esto se comparó como cambia el ajuste de los modelos y las significancias de las variables al buscar explicar los accidentes en otro periodo.

Se presentan a continuación el resumen del indicador R^2 ajustado, el total de accidentes de cada periodo y el número de intersecciones de cada modelo según sus formas conceptuales en los seis periodos. Además, se contabilizó en cada periodo el número de intersecciones que presentaban cero accidentes en el periodo de análisis.

Tabla 6.2 R^2 de los modelos más adecuados en cada uno de los periodos de análisis

Formas Conceptuales		AM	PMD	PT	FP	DIA COMPLETO	DIURNO
R²	1	28.44	22.29	19.04	40.17	42.23	40.91
	2	28.32	22.03	18.92	40.61	42.33	40.89
	3	28.57	21.99	18.80	40.37	41.78	40.55
	4	28.57	22.29	19.06	40.48	42.11	40.81
	5	29.84	23.32	19.07	39.87	42.22	40.87
	6	22.89	15.85	13.14	28.99	27.07	27.56
	7	24.57	24.40	17.64	33.52	40.01	37.65
	Promedio	27.32	21.74	17.95	37.72	39.68	38.46
Nº Accidentes		372	449	520	1,635	3,064	2,816
Nº Intersecciones		271	272	272	270	268	268
Nº Inter sin Accidentes		100	99	81	11	6	7

El periodo que presenta mejores ajustes en los modelos es el día completo. No obstante, como se observa en la Tabla 6.2 el modelo 6 (que no considera variables de flujo), logra el mayor ajuste en el periodo fuera de punta, lo cual se condice con lo esperado ya que en este periodo el flujo vehicular disminuye y por consiguiente, tiene una menor importancia en su efecto con la ocurrencia de accidentes.

Se obtuvo como resultado que en los periodos FP, Día y Día Completo los ajustes de los modelos en general son mejores pasando de órdenes de 25% a 40% de R^2 . Por otro lado, sucede que mientras menor es el número de intersecciones con cero accidentes mejor es la estimación.

Ahora bien, si solo se comparan los periodos puntas entre sí, se observa que a pesar que la cantidad de accidentes es mayor en el periodo punta tarde, y la cantidad de intersecciones con cero accidentes es menor, es justamente este periodo en donde los modelos poseen los menores ajustes.

Finalmente, en el Anexo 6.1 se presenta un resumen de cada uno de los siete modelos, donde se comparan los valores y la significancia estadística de los coeficientes asociados a cada variable.

6.2.3 Modelo de regresión lineal múltiple final

En esta sección se presentan los resultados del mejor modelo estimado luego de haber realizado la depuración de la base de datos a través de la eliminación de todas aquellas intersecciones con residuos atípicos (según lo que se explicó en la sección 6.1.3). Con esto se buscó disminuir las distorsiones generadas por valores en donde las diferencias entre lo observado y lo modelado son muy grandes.

Como resultado de este proceso y tomando como base las formas conceptuales analizadas anteriormente, se obtuvo que al replicar las mejores combinaciones de variables (modelos 1, 2, 3, 4 y 5) y buscando maximizar el ajuste del modelo y la significancia de las variables, se obtuvo la forma conceptual final mostrada en la Tabla 6.4.

Tabla 6.3 Resultados de la forma conceptual más adecuada, Modelo 8

Variables	Coefficiente	Test t	Valor p
Constante	-10.934	-5.732	0.000
N°Pistas RS	1.291	2.779	0.006
Angulo Cat	0.036	1.744	0.082
Densidad Peatonal	2.317	5.123	0.000
N°Cruces Peatonales	0.670	2.545	0.012
TMD Intersección	0.0002	5.577	0.000
TMD Camiones	0.0008	2.282	0.023
TMD Taxi buses	0.0004	1.702	0.089
R2 ajustado	46.56		

Al eliminar los datos con valores de residuos atípicos se mejora en alrededor de 5% la capacidad de predicción del modelo obteniendo el resultado mostrado. En este modelo las variables de mayor peso son el tráfico total de la intersección, la densidad peatonal y la constante. Además, están las variables del número de pistas en la rama secundaria, el número de cruces peatonales y el flujo de camiones. Luego y solo si se considera un intervalo de confianza del 90%, las variables ángulo de la intersección y flujo de taxibuses resultan significativos estadísticamente.

6.2.4 Modelo de regresión Poisson final

En los modelos estimados del tipo Poisson, el método estadístico de base es la máxima verosimilitud. En el *software* utilizado se obtiene como resultado un porcentaje de explicación de la desviación del modelo, el cual es comparable con el R^2 ajustado en los modelos de regresiones lineales. A continuación, se presentan en la Tabla 6.5 los resultados de la forma conceptual del modelo 8 (sección 6.2.3).

Tabla 6.4 Resultados de la modelación Poisson, con la forma conceptual del modelo 8 como base

Variables	Coefficiente	Chi-Cuadrada	Valor p
Constante	0.7398		
N°Pistas RS	0.1143	16.994	0.0000
Angulo Cat	0.0049	14.079	0.0002
Densidad Peatonal	-0.4323 (valor igual a 1) -0.2522 (valor igual a 2)	63.768	0.0000
N°Cruces Peatonales	0.0673	18.945	0.0000
TMD Intersección	0.00002	113.410	0.0000
TMD Camiones	0.00004	6.529	0.0106
TMD Taxi buses	0.00004	9.663	0.0019
R2 ajustado		45.96	
Desviación Modelo		618.18	
Desviación Residuo		687.57	

Al comparar estos resultados con la versión del modelo 8 (regresión lineal), no se observa una mejoría apreciable del nivel de ajuste. Un aspecto a destacar es el signo de la variable densidad peatonal, la cual fue considerada como categórica en esta modelación. En otras palabras, los coeficientes asociados al nivel 1 y 2 de la variable se calcularon con respecto al del nivel 3, es decir, a la máxima densidad peatonal definida, es por esto que en las intersecciones con menor valor a 3 el coeficiente es negativo, lo cual se condice con lo esperado y los resultados anteriores, pues a menor nivel de densidad peatonal, menor es el nivel de accidentes.

Considerando como base el modelo 8 se buscó simplificar y combinar otras variables para revisar hasta qué nivel la regresión Poisson puede predecir los accidentes.

Como se observa en la Tabla 6.5, el valor p más alto es de la variable de flujo de camiones, luego se realiza la modelación sin esta variable y se obtienen los siguientes resultados de los parámetros de ajuste.

Tabla 6.5 Ajustes del modelo excluyendo la variable de flujo de camiones

Parámetro	Valor
R2 ajustado	45.62
Desviación Modelo	611.65
Desviación Residuo	694.09

El detalle de los valores p y los coeficientes de cada variable se encuentran en el Anexo 6.2. De los nuevos valores p asociados a las variables resultan todos bajo 0.0003 a excepción del valor p del flujo de taxibuses (valor p = 0,0113). Por consiguiente, se realiza el mismo proceso y se modeló eliminando ahora la variable de tráfico de Taxibuses. Se obtuvo los resultados mostrados en la Tabla 6.7. En mayor detalle se puede revisar los resultados en el Anexo 6.2.

Tabla 6.6 Ajustes del modelo excluyendo el flujo de camiones y taxibuses

Parámetro	Valor
R2 ajustado	45.27
Desviación Modelo	605.23
Desviación Residuo	700.51

En este caso el modelo se ha simplificado manteniendo solo cinco variables de las siete iniciales y en cuanto a su capacidad predictiva solo ha disminuido de 45,96 a 45,27. Por lo cual el modelo simplificado resulta competente y adecuado en términos generales.

Además de esta forma conceptual se probaron algunas variantes, intentando iterar de tal manera de observar los ajustes incluyendo variables como la Visibilidad y el Número de Paraderos. No obstante, no se obtuvo un modelo con mejores ajustes estadísticos que el modelo presentado. Los resultados de la modelación y de las pruebas mencionadas se encuentran en el Anexo 6.2.

6.2.5 Discusión de resultados

En la modelación lineal se obtuvo que de las variables de infraestructura, el número de pistas en la rama secundaria resultó como la variable que más se repitió en los modelos probados resultando siempre significativa estadísticamente y también se obtuvo siempre un valor de su coeficiente asociado positivo, lo cual resulta lógico, pues a mayor número de pistas, mayor es el riesgo de accidentes. Por otra parte, la variable ángulo de la intersección (categorizada en 6 rangos) resultó ser una variable importante, la cual resultó significativa en todos los cinco modelos completos que se analizaron (Modelo 1, 2, 3, 4 y 5). No obstante, de los resultados se tuvo que a medida que el ángulo se acercaba a 90° aumentaba el valor de accidentes, esto pues el coeficiente siempre resultó positivo. Esto implica que en aquellas intersecciones perpendiculares existen más accidentes que en aquellas más anguladas. En general se pueden conjeturar dos hipótesis. La primera sería pensar que esto se debe a que al existir muchas más intersecciones con 90° de ángulo hay mayor probabilidad de tener intersecciones con numerosos accidentes. La segunda hipótesis apuntaría a que en aquellas intersecciones donde el ángulo es menor resultaron ser en su mayoría intersecciones con menor demanda de flujo y de sectores más periféricos y, por consiguiente, presentaban menores índices de accidentabilidad.

En lo que respectan las variables de operación se obtuvo que la variable de densidad peatonal siempre resultó significativa, lo cual es un indicio de la importancia que genera el efecto de la presencia de peatones en una intersección. La variable visibilidad categorizada (en 6 rangos según la distancia de visibilidad calculada), si bien no resultó significativa al 90% de confianza, si se considera un 80% de confianza, sí es significativa. Sumado a esto los ajustes de los modelos donde se considera la visibilidad mejoran y según la literatura la visibilidad es una variable importante en los modelos de accidentes (Salifu, 2003, Wang *et. al.*, 2011, entre otros). El aspecto a destacar en este sentido es que al modelar con la visibilidad como valor continuo no mejora su significancia ni el ajuste en el modelo, es decir, esta variable para el nivel de predicción de este tipo de modelos se puede trabajar en rangos categorizando por distancia de visibilidad. Sumado a esto se probaron modelos agregando variables como el número de paraderos y el número de cruces peatonales. En estos, si bien el ajuste se mantuvo, dichas variables no resultaron significativas a un 80% de confianza, por lo cual pueden no ser consideradas.

En lo que respecta las variables asociadas al flujo se obtuvo, que son las que mayor explican la ocurrencia de accidentes y al probarlas en los modelos se observó que el valor del flujo total en la intersección, el flujo de vehículos livianos, el flujo de camiones y el flujo de taxibuses resultaron ser variables estadísticamente significativas, además en todos los casos se obtuvo elevados valores del test t asociados a estas variables. Y esto se refuerza con el ejercicio de la modelación de las formas conceptuales 6 y 7, en donde solo utilizando variables de flujo se alcanzaba un 38% de ajuste de R^2 lo cual es comparativamente igual a los ajustes de los modelos completos (modelos 1, 2, 3, 4 y 5).

En lo que se observó al replicar los modelos en otros periodos, se obtuvo que el ajuste de los modelos guarda relación directa con el número de accidentes en las bases de datos, pues donde mayor era éste, mejor era el ajuste de R^2 . Asimismo, cuando las bases de datos de accidentes contienen muchas intersecciones con cero accidentes los modelos lineales de regresión pierden considerablemente su ajuste de R^2 y las significancias estadísticas de sus variables. No obstante, resulta importante destacar que el periodo punta tarde a pesar de tener mayor número de accidentes y menor cantidad de intersecciones sin accidentes que los periodos punta mañana y punta medio día es en donde todos los modelos estimados poseen los menores ajustes. Es entonces necesario de variables más precisas para este periodo puntual, o de otro tipo de factores que pudieran escapar de los alcances de las variables medidas para poder explicar los accidentes.

Se observa que al utilizar la regresión del tipo Poisson, el ajuste de los modelos se mantiene en el mismo orden de magnitud del valor R^2 . No obstante, es importante destacar que al ser una distribución más *ad-hoc* a la de los accidentes resulta teóricamente más correcto utilizar esta herramienta. Es por esto que, al disminuir la cantidad de variables a cinco, el modelo mantuvo su capacidad predictiva en el orden 45% de R^2 .

6.3 Modelación no lineal

En esta sección se presentan los resultados de la modelación no lineal realizada. Se probaron formas funcionales no lineales, tales como polinomiales y funciones exponenciales. Estas últimas son ampliamente utilizadas en la literatura (Cunto *et. al.* (2014), Salifu (2003), Vieira Gomes (2013) entre otros). Por ello y en base a los resultados que se obtuvieron se realizó un análisis con mayor profundidad utilizando como estructura base para los modelos la siguiente forma funcional de la Ecuación 6.2.

$$\text{N}^\circ \text{Accidentes} = \beta \times \text{Flujo vehicular} \times \exp(\Sigma \text{Var Infra} \cdot \theta + \text{Var Operación} \cdot \alpha + \text{Var Entorno} \cdot \gamma) \quad (6.2)$$

Donde β es un coeficiente estimado asociado a la variable de flujo y θ, α, γ los coeficientes asociados a las variables de infraestructura, operación y entorno de la intersección.

En base a este modelo se realizaron ciertas pruebas apuntando a analizar cómo variaban los ajustes. Para ello se probó agregar ciertos parámetros de potencias asociándolos a algunas variables, así como la combinación de distintas variables independientes.

En general, al estimar los modelos se buscó minimizar la cantidad de variables buscando simplificarlos. Las formas funcionales que se presentan en la Tabla 6.7 fueron aquellas donde se obtuvo los mejores resultados de ajuste R^2 .

De la Tabla 6.7 y de las pruebas realizadas, se concluye que los modelos con menos de cinco variables resultan insuficientes para lograr una estimación adecuada del número de accidentes, pues como se observa en los modelos 6, 18 y 9 se obtienen bajos niveles de ajustes. Asimismo se observa la importancia de los parámetros estimados para adecuar mejor el modelo y cada variable. En este caso es destacable la diferencia que existe entre el modelo 6 y 7 en donde a la variable TMD Intersección se consideró con un coeficiente asociado en el modelo y un parámetro de potencia en el modelo 7, mientras que en el modelo 6 solo se consideró un parámetro de ajuste de potencia, el resto de la forma funcional es igual en ambos modelos no obstante, la diferencia es importante obteniéndose un 31,7% de ajuste en el modelo 6 y un 50,5% en el 7.

Tabla 6.7 Formas conceptuales y ajustes de los modelos no lineales analizados

ID	Forma Conceptual del Modelo y parámetros estimados	R ² Ajustado
1	N° Accidentes = a TMD_INTER • EXP (b N° PISTAS_RS + c VISIBILIDAD ^z + d DENSIDAD PEATONAL + e N° CRUCES PEATONALES)	50.33
2	N°_ACC = a TMD_INTER ^z • EXP (b N° PISTAS RS + c VISIBILIDAD + d DENSIDAD PEATONAL + e N° CRUCES PEATONALES)	50.53
3	N° ACC = a TMD_INTER • EXP (b N° PISTAS RS + c ANGULO_Cat ^z + d DENSIDAD PEATONAL + e N° CRUCES PEATONALES)	51.64
4	N° ACC = a TMD_INTER ^z • EXP (c N° PISTAS RS + d ANGULO_Cat + e DENSIDAD PEATONAL + f N° CRUCES PEATONALES)	51.03
5	N° ACC = a TMD_INTER ^z + c N° PISTAS RS ^w + d DENSIDAD PEATONAL ⁿ + e N° CRUCES PEATONALES ^m + f ANGULO_Cat ^y	43.07
6	N° ACC = TMD_INTER ^z • EXP (b N° PISTAS RS + c DENSIDAD PEATONAL + d N° PARADEROS + g N° CRUCES PEATONALES)	31.70
7	N° ACC = a TMD_INTER ^z • EXP (b N° PISTAS RS + c DENSIDAD PEATONAL + d N° PARADEROS + g N° CRUCES PEATONALES)	50.51
8	N° ACC = a TMD_INTER • EXP (b N° PISTAS RS + c ANGULO_Cat ^z + d DENSIDAD PEATONAL + e VISIBILIDAD)	50.70
9	N° ACC = a TMD_INTER ^z • EXP (b N° PISTAS RS + c DENSIDAD PEATONAL)	36.34
10	N° ACC = a TMD_INTER ^z • EXP (b N° PISTAS RS + c DENSIDAD PEATONAL + g N° CRUCES PEATONALES)	50.45
11	N° ACC = a TMD_INTER • EXP (b N° PISTAS RS + c DENSIDAD PEATONAL + d N° PARADEROS + e N° CRUCES PEATONALES + VISIBILIDAD ^z)	50.47
12	N° ACC = a TMD_INTER • EXP (b DENSIDAD PEATONAL + c N° CRUCES PEATONALES + d N° PARADEROS + VISIBILIDAD ^z)	48.35
13	N° ACC = a TMD_INTER • EXP (b DENSIDAD PEATONAL + c N° CRUCES PEATONALES + d N° PARADEROS + e N° PISTAS RS)	50.49
14	N° ACC = a TMD_INTER • EXP (b DENSIDAD PEATONAL + c N° CRUCES PEATONALES + e N° PISTAS RS)	50.52
15	N° ACC = a TMD_INTER ^z • EXP (b N° PISTAS RS + c DENSIDAD PEATONAL + d N° PARADEROS + e Pista Izquierda RP + f Mediana RP + g N° CRUCES PEATONALES)	50.50
16	N° ACC = a TMD_INTER ^w • b TMD_CAM_DIA ^z • EXP (c N° PISTAS RS + d ANGULO_Cat + e DENSIDAD PEATONAL + f N° CRUCES PEATONALES)	51.06
17	N° ACC = a TMD_INTER ^w • b (TMDRS/TMDRP) ^z • EXP (c N° PISTAS RS + d ANGULO_Cat + e DENSIDAD PEATONAL + f N° CRUCES PEATONALES)	51.50
18	N° ACC = a + b (TMDRS/TMDRP) ^z • EXP (c N° PISTAS RS + d ANGULO_Cat + e DENSIDAD PEATONAL + f N° CRUCES PEATONALES)	18.26

Nota: a, b, c, d, e, f, g son coeficientes estimados de multiplicación asociados a cada variable y n, m, y, w, z son parámetros estimados de potencia asociados a cada variable.

El mejor modelo estimado fue el 3. Al comparar los resultados de la Tabla 6.7, el modelo 3 con el modelo 4 a pesar de ser prácticamente iguales, se observa que en este último el ajuste es menor lo cual implica que el considerar la variable Angulo como no lineal en su aporte al número de accidente resulta adecuado. En contraposición al realizar la comparación para la variable de Visibilidad en los modelos 1 y 2 se encuentra un resultado diferente en donde el modelo en el cual no se asocia un parámetro de potencia a la variable resulta con un mejor ajuste general de R^2 . Ambos resultados de las variables Angulo y Visibilidad se refuerzan en los resultados del modelo 8 en donde su ajuste es superior al considerar las respectivas formas adecuadas de cada variable en el modelo.

Otro aspecto a mencionar es el caso del modelo 13 en donde al simplificarlo, quitando la variable Número de Paraderos, mejora el ajuste del modelo (ver modelo 14). Por otro lado en el modelo 15, a pesar de combinar y utilizar siete variables, el ajuste solo mejoró marginalmente, luego es preferible utilizar el modelo más sencillo.

Finalmente, al incorporar otra variable de flujo en los modelos se observan elevados índices de ajustes con respecto a los casos analizados, como lo son los modelos 16 y 17 en donde destaca el modelo 17 con un R^2 de 51,5 al incluir el efecto de la variable TMDRS/TMDRP la cual relativiza las diferencias de flujo entre la rama principal y la secundaria. No obstante, al realizar la modelación únicamente con esta variable, es decir, manteniendo la forma conceptual del modelo 17 intercambiando la variable TMD Intersección por la de TMDRS/TMDRP el ajuste disminuye de manera importante alcanzando solo un 18%.

Como síntesis de lo analizado y observado en los resultados el modelo 3 fue en el que se obtuvo un mejor ajuste de R^2 alcanzando un 51.6% con solo cinco variables.

6.4 Validación de modelos estimados

Para realizar la validación se contrastaron los resultados de los modelos con respecto a los datos observados. En este caso para realizar dicho análisis, se utilizó una porción de la base de datos, es decir, un 20% de las intersecciones. Con éstas se realizó la validación del modelo 3 descrito y analizado en la sección 6.3.

En este caso el 20% de las intersecciones consideradas fueron escogidas aleatoriamente de la base de datos. Se realizó el análisis tanto para los modelos de regresión lineal múltiple, regresión Poisson y el modelo no lineal, los cuales fueron escogidos como los más adecuados.

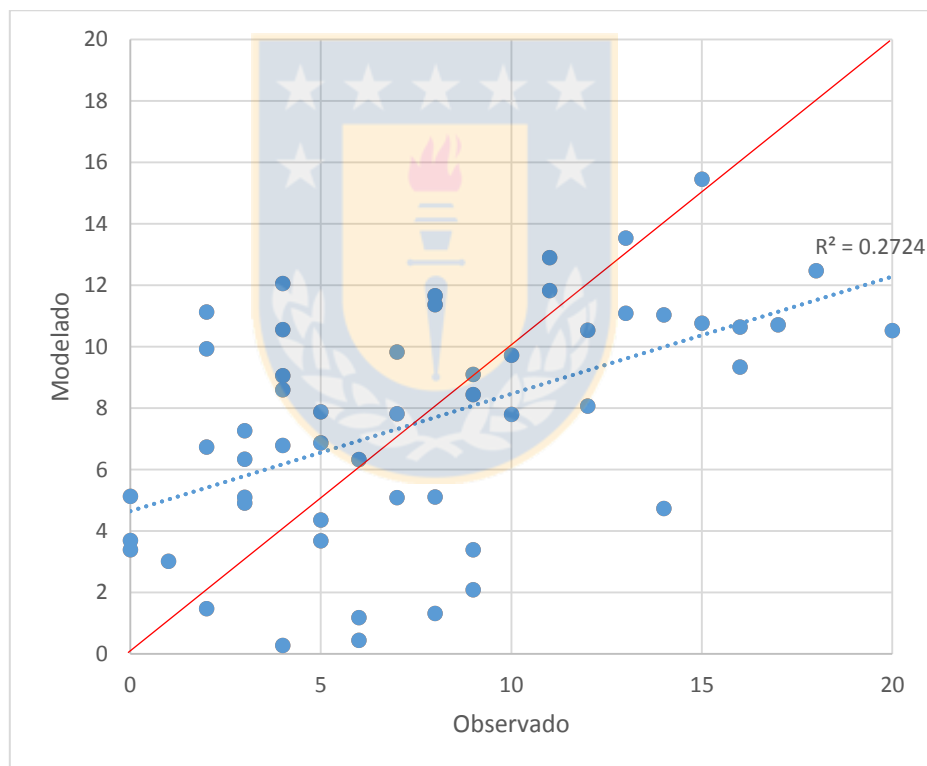


Figura 6.2 Número de accidentes modelados versus observados con el modelo de regresión lineal

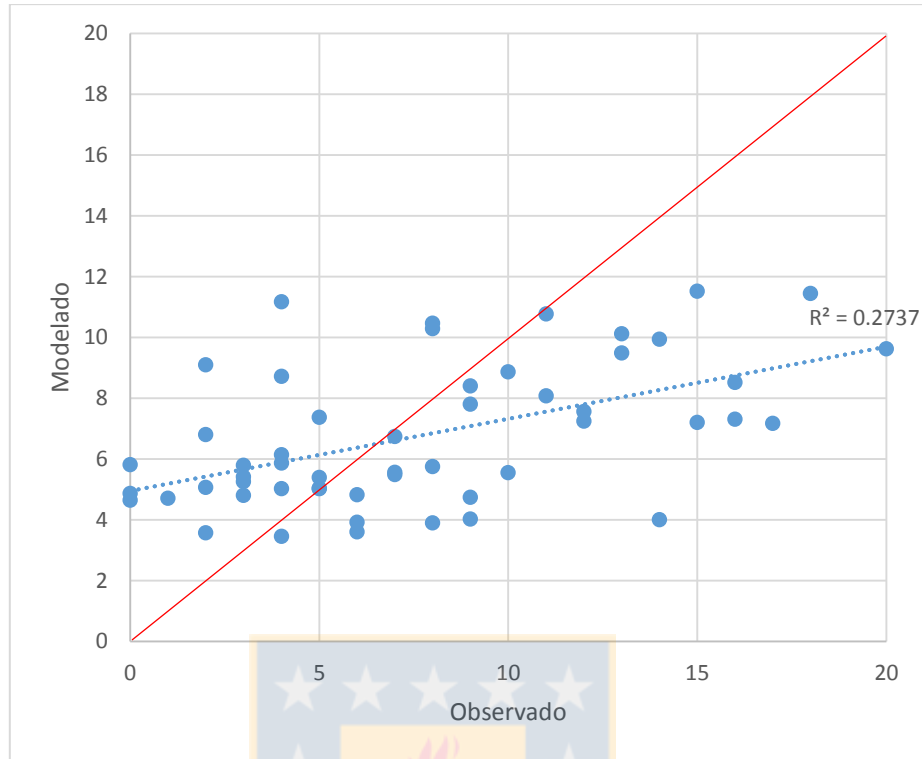


Figura 6.3 Número de accidentes modelados versus observados con el modelo de Poisson

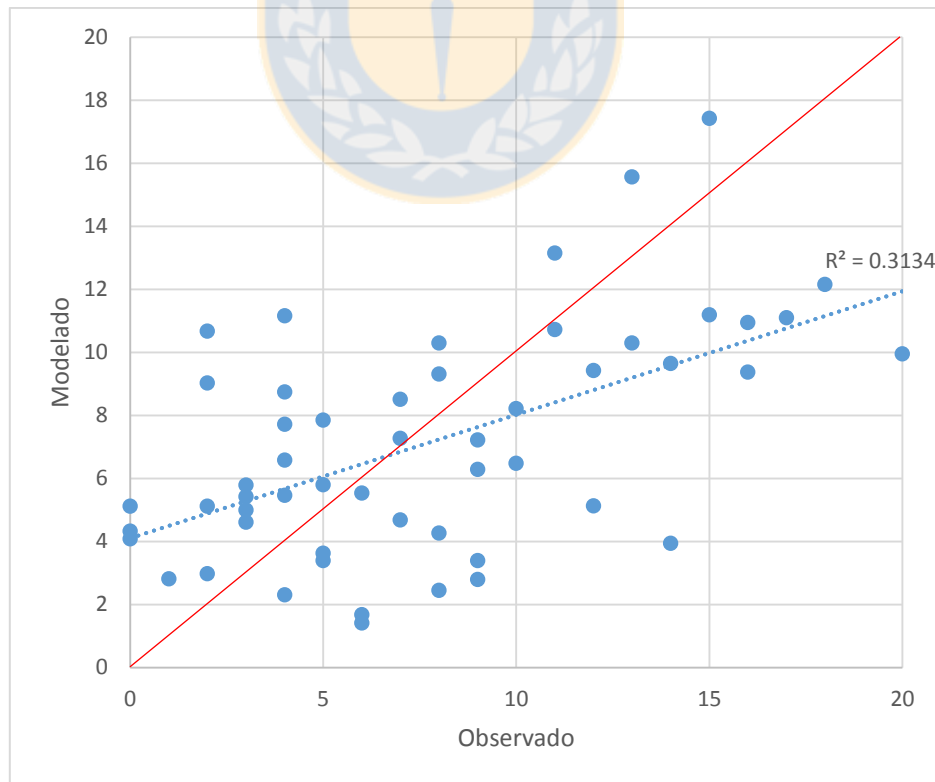


Figura 6.4 Número de accidentes modelados versus observados con el modelo no lineal

Como se puede observar en las Figuras 6.2, 6.3 y 6.4 los ajustes en la validación son muy bajos. Esto se debe a la baja cantidad de datos para validar, asimismo, se comprueba la capacidad predictiva limitada de dichos modelos, no obstante entre sí se observa que el modelo no lineal replica de mejor manera los accidentes en las intersecciones que los modelos lineal y Poisson, esto pues gráficamente se puede observar menor dispersión en los datos y no se observa ninguna tendencia explícita que pueda generar cierto sesgo en la predicción.

Además, se presenta en las siguientes figuras un análisis por zona de las diferencias obtenidas entre los resultados del número de accidentes según el modelo calibrado, versus el número de accidentes observado.

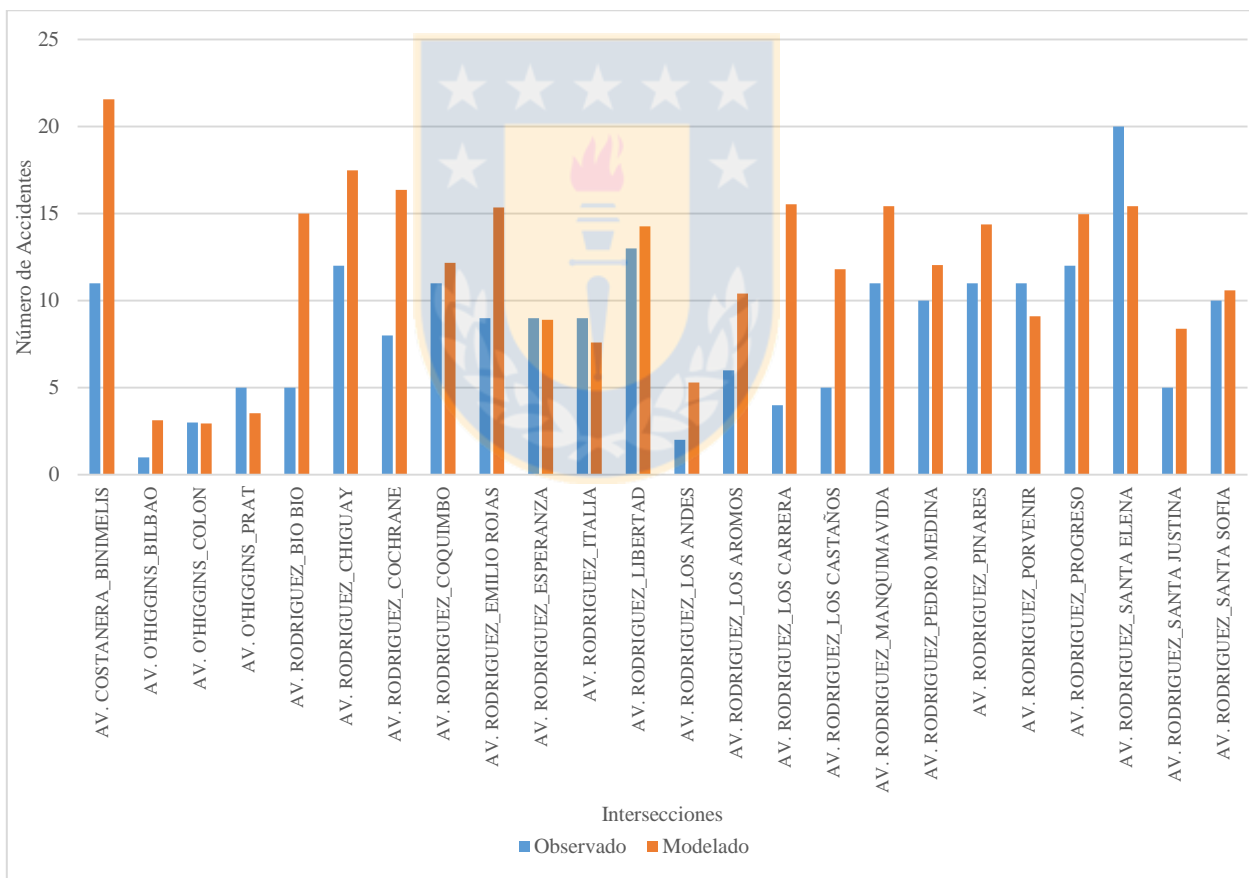


Figura 6.5 Número de Accidentes en intersecciones de Chiguayante

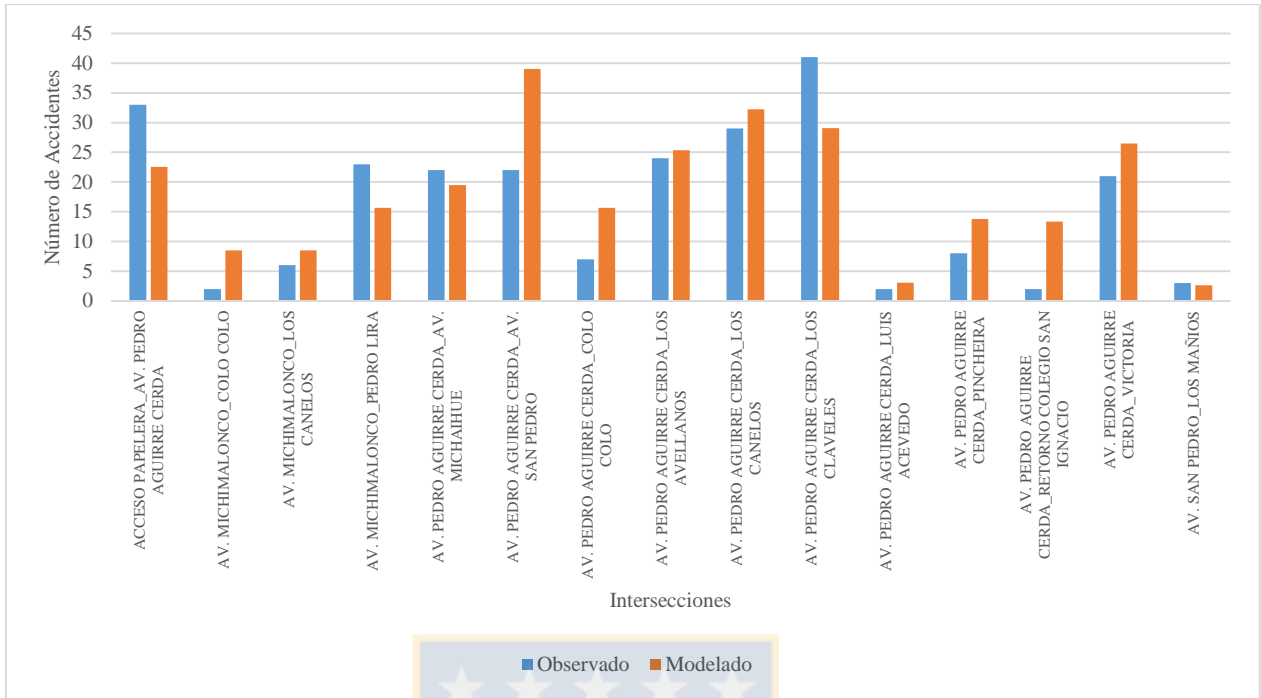


Figura 6.6 Número de Accidentes en intersecciones de San Pedro de la Paz

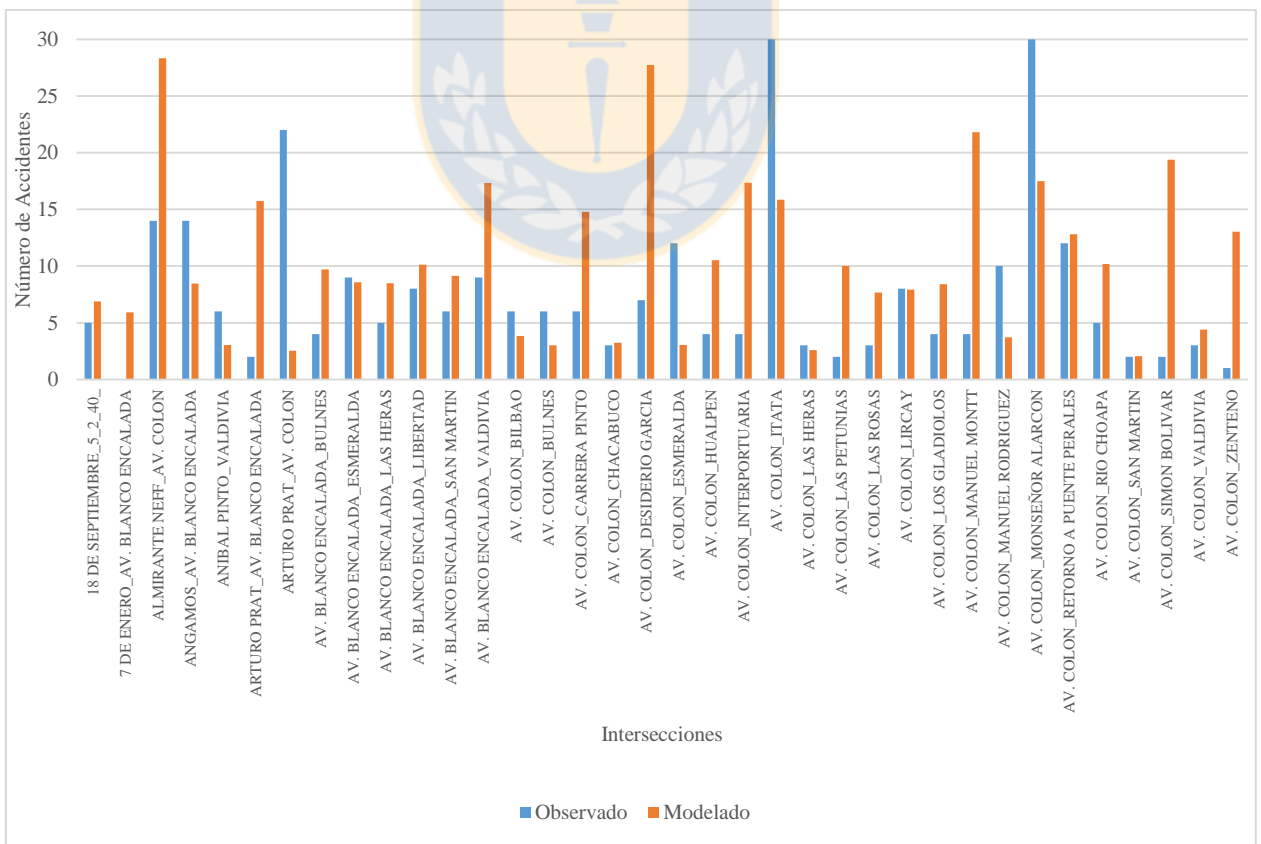


Figura 6.7 Número de Accidentes en intersecciones de Talcahuano

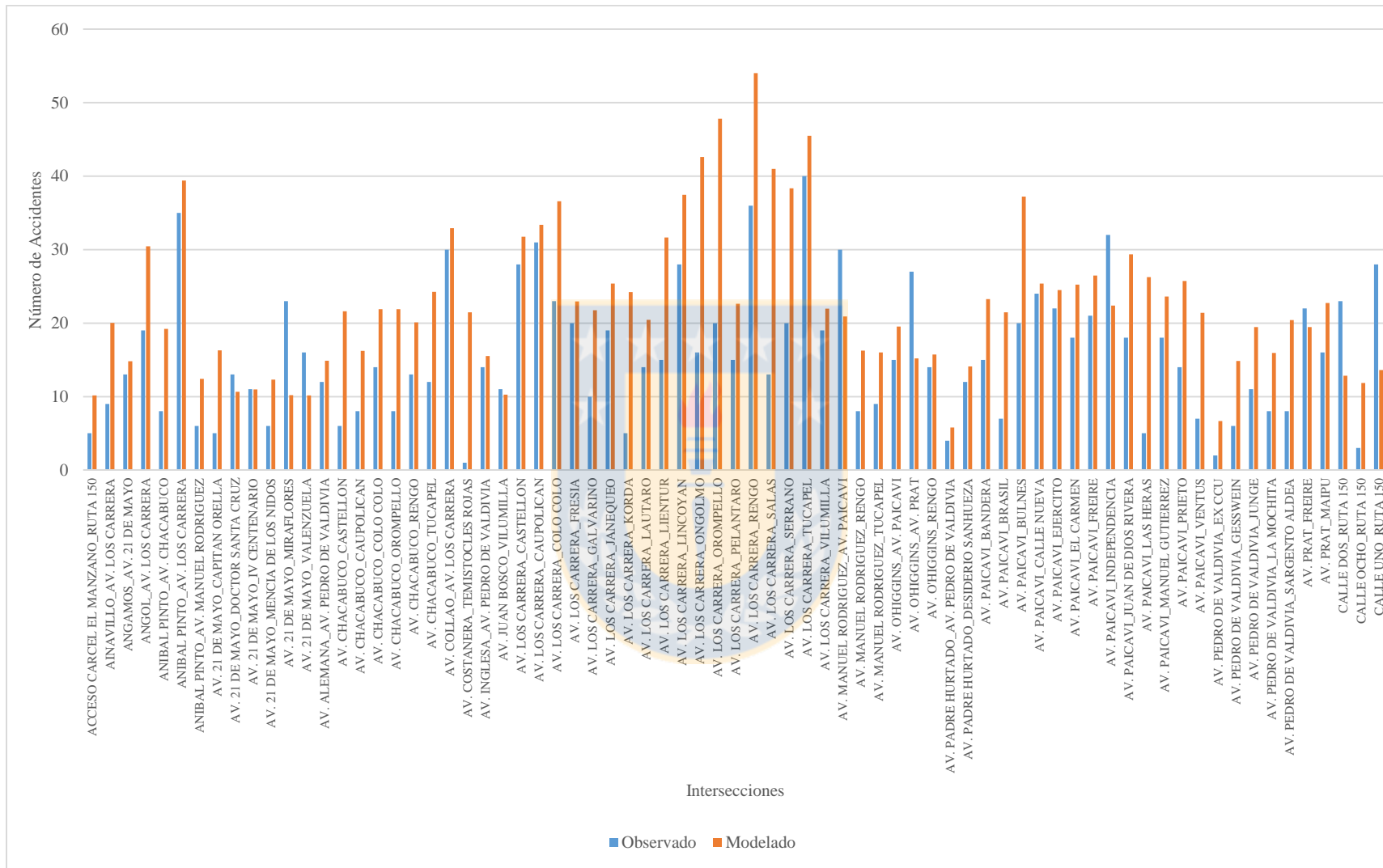


Figura 6.8 Número de Accidentes en intersecciones de Concepción con flujo mayor a 30.000 (veq/día)

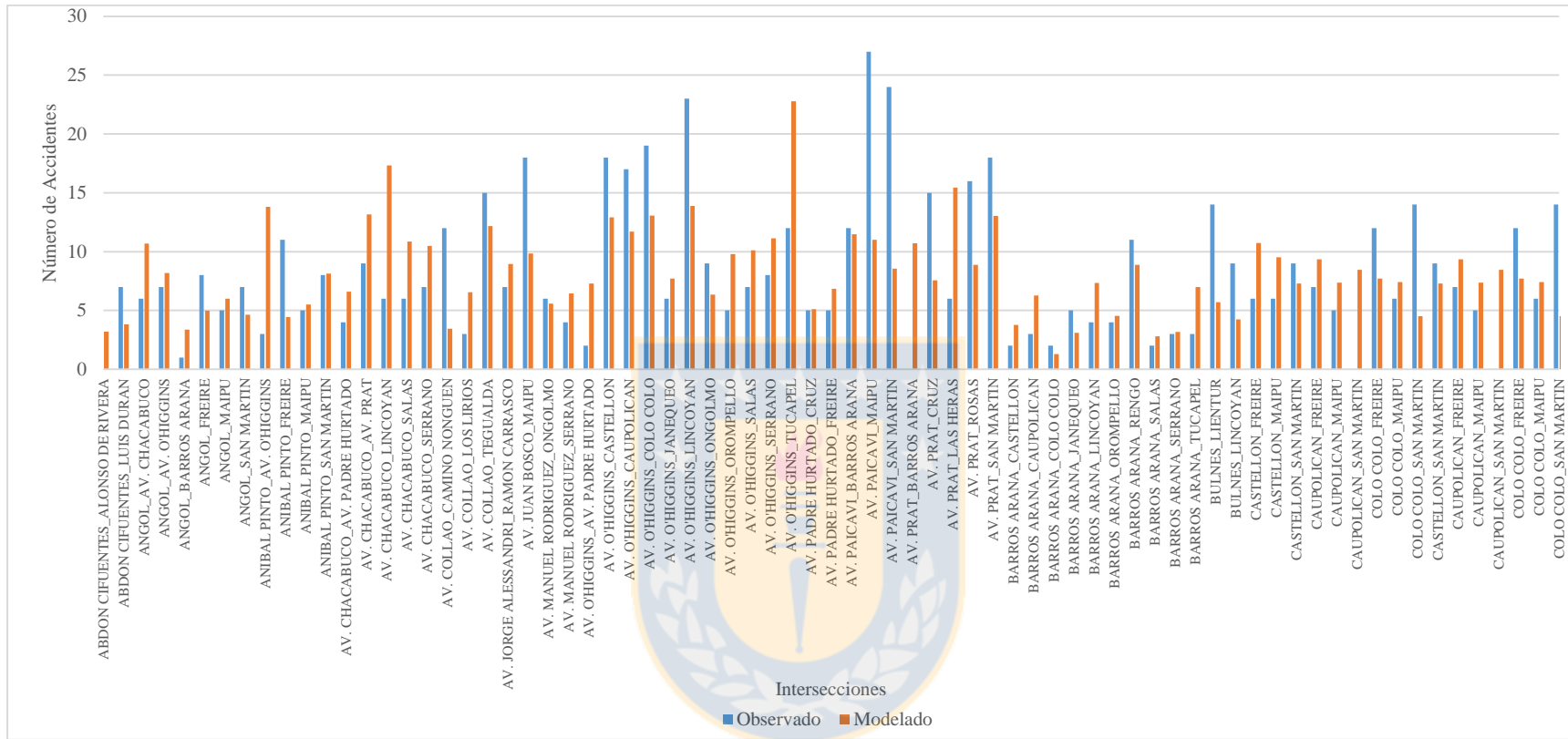


Figura 6.9 Número de Accidentes en intersecciones de Concepción con flujo menor a 30.000 (veq/día)

De las Figuras 6.5, 6.6, 6.7, 6.8 y 6.9, se desprende una relación entre la capacidad predictiva del modelo, la cual disminuye en las intersecciones donde existe una elevada diferencia entre los flujos de ambas calles. Es decir, en intersecciones donde una de las vías es de carácter primario y de alto flujo y la segunda vía es de carácter local y de bajo flujo se observa a que las diferencias entre lo predicho y lo observado son mayores. Además, se observa que en intersecciones donde existe bajo flujo vehicular total (Figura 6.9) y en aquellas de carácter céntrico, donde transita transporte público y además peatones, se generan también elevadas diferencias entre lo estimado por el modelo y el número de accidentes, en donde se ve que el modelo subestima la cantidad de accidentes ocurridos.

En la Tabla 6.8 se presenta el resumen de los indicadores de estas comparaciones mostradas en las Figuras 6.5, 6.6, 6.7, 6.8 y 6.9.

Tabla 6.8 Indicadores de comparación en la estimación de los accidentes con el modelo no lineal versus los datos observados.

Item	Diferencias			Accidentes		
	Mínima	Maxima	Media	Mínima	Maxima	Media
San Pedro de la Paz	0.35	17.03	6.36	2	41	16.33
Talcahuano	0.06	20.71	7.10	0	30	7.52
Chiguayante	0.06	11.54	4.07	1	20	8.45
Concepción intersecciones con flujo > 30.000	0.04	27.98	9.21	1	40	15.57
Concepción intersecciones con flujo < 30.000	0.11	15.99	4.41	0	27	8.27

Como se puede observar el modelo estimó con el menor grado de exactitud el número de accidentes en las intersecciones de la comuna de Talcahuano, y luego en aquellas intersecciones de Concepción de elevado flujo. Aquí se puede observar a priori que los resultados del modelo funcionan de mejor manera en intersecciones con niveles intermedios de flujo vehicular, como es el caso de intersecciones de Chiguayante, San Pedro de la Paz, y Concepción con flujo menor a 30.000 vehículos por día.

6.5 Comparación de modelos en la literatura

En este caso se realizó una comparación de una validación de un modelo revisado en la literatura, de modo tal de contrastar los resultados con el caso de estudio. En la Figura 6.10 se presenta la validación del modelo calibrado por Cunto *et. al.* (2014), donde en su caso también existe una gran dispersión de los datos, no obstante se observa una diferencia con respecto a lo presentado en la sección 6.4, y radica en que existe una clara tendencia del modelo a subestimar el número de accidentes cuando el valor de éstos son elevados, mientras que para intersecciones con menores niveles de accidentes se observa una tendencia más leve a sobre estimar los accidentes.

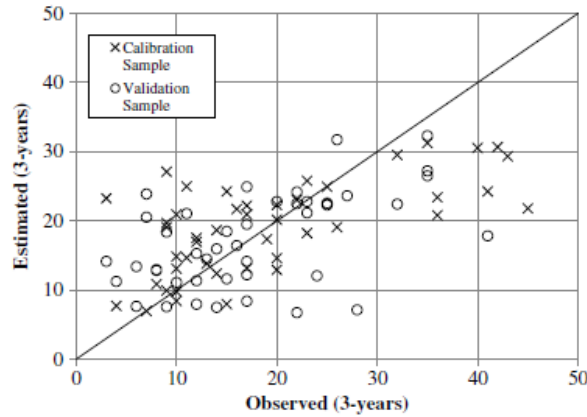


Figura 6.10 Número de accidentes modelados y observados en el modelo de Cunto *et.al.* (2014)

Otro aspecto que se analizó a modo de ejemplo fue cómo replicaban los datos de accidentes de Concepción otros modelos externos de la literatura. Para ello se utilizó de la información disponible del modelo de Vieira Gomes (2013) en donde utilizando los coeficientes y la forma conceptual de dicho estudio se replicó y calcularon los números de accidentes en las intersecciones con las cuales se realizó la validación.

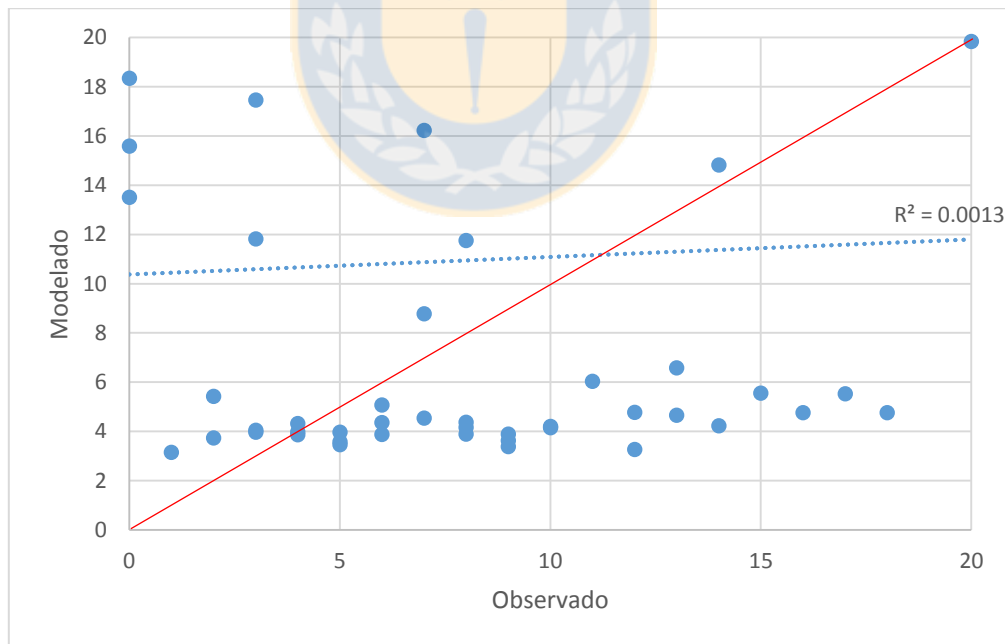


Figura 6.11 Número de accidentes estimados con el modelo de Vieira Gomes (2013) versus el número de accidentes observados

Además de esto, se realizó la comparación entre los resultados de número de accidentes estimado para el Gran Concepción con los parámetros y el modelo de Vieira Gomes (2013) versus el modelo 3 no lineal estimado en el presente estudio (ver Tabla 6.8). Dichos resultados se presentan en la Figura 6.7.

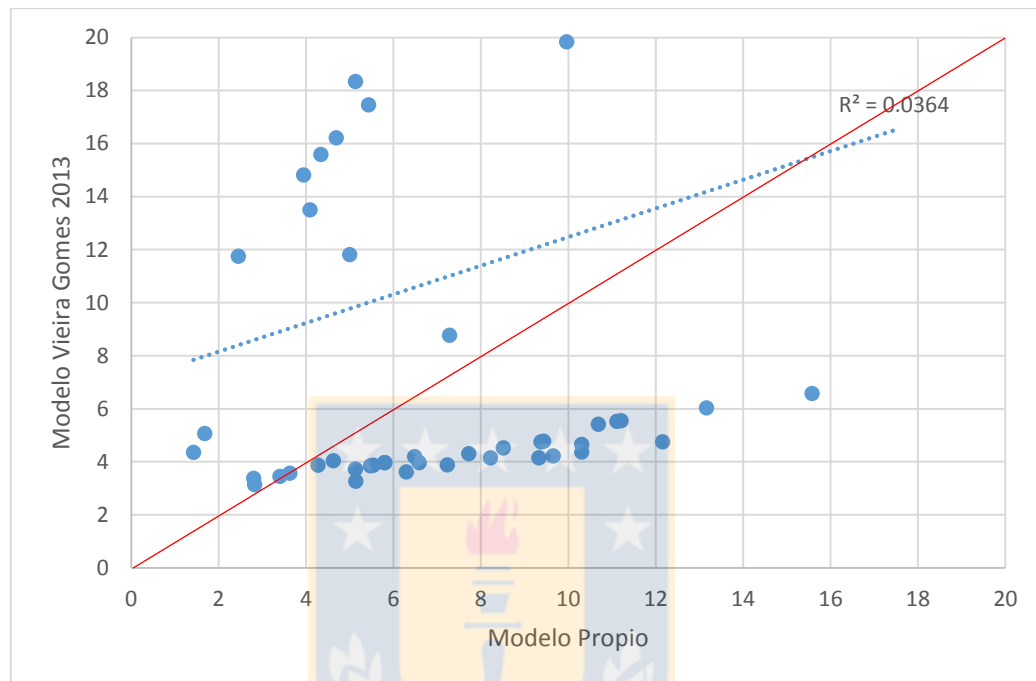


Figura 6.12 Número de accidentes según el modelo de Vieira Gomes (2013) versus según el Modelo Propio

Como se puede observar no existe concordancia alguna entre las estimaciones y los ajustes no son adecuados. Esto corrobora lo revisado en la literatura sobre la replicación y extrapolación de los modelos de accidentes. Pues como se observa en este ejemplo este modelo no es extrapolable al caso del Gran Concepción, ya que se observan estimaciones y resultados radicalmente distintos. Por ejemplo existen puntos donde con el modelo de Vieira Gomes se obtienen valores de accidentes entre 3 y 6. En cambio, con el modelo propio se estiman en estas mismas intersecciones, valores de accidentes que van desde los 3 hasta 14 accidentes.

Es importante destacar la importancia de las condiciones que no son propias de las intersecciones y que radican en aspectos endógenos de cada situación y contexto en el que ocurran los accidentes, por consiguiente, esta comparación de los modelos no resulta adecuada.

6.6 Conclusiones

En este capítulo se estimaron modelos tanto lineales como no lineales, de éstos, se logró seleccionar un modelo con un mayor nivel de explicación (de cada tipo de modelación, Regresión Lineal, Poisson y No Lineal), los cuáles fueron validados y comparados con un modelo de predicción de accidentes existente en la literatura.

Se observó que las variables relacionadas con la operación tienen una mayor incidencia en la predicción de accidentes que las variables relacionadas con la infraestructura (alcanzando por sí solas un ajuste de 37%). Esto se ve reflejado en que las variables de operación son siempre significativas y con elevados test t. Un ejemplo de esto es que en todos los modelos estimados se obtuvo como resultado valores de test t superiores a 3 para la variable de flujo en la intersección. En contraste a esto, las únicas variables de infraestructura que resultaron significativas fueron el número de pistas de la rama secundaria y el ángulo de la intersección, las que además, presentaron test t bajos (inferiores a 2,1).

De los modelos lineales y no lineales con los mejores ajustes se obtuvieron valores de R^2 de 46% y 51%, respectivamente. Si bien no hay grandes diferencias en el ajuste de ambos, se logró una mejoría en el modelo no lineal, ya que se simplificó en cuanto al número de variables utilizadas, (disminuyendo de siete a cinco variables).

El modelo estimado con regresión lineal de Poisson presentó un ajuste de 45%, el cual a pesar de ser inferior a los otros modelos estimados, se hizo con una menor cantidad de variables que el modelo lineal llegando a una simplificación igual a la del modelo no lineal, es decir, utilizándose cinco variables.

Aun cuando los modelos presentan diferencias en su planteamiento y resultados obtenidos, otro aspecto a destacar en todas las regresiones lineales hechas es que los test t de las constantes siempre resultaron elevados en relación a aquellos de las demás variables. Esto guarda relación con los aspectos que no están siendo captados por las variables utilizadas en los modelos. Luego resulta

importante estar conscientes de los alcances y aspectos que no se están representando en la modelación.

Siguiendo con esto resulta importante concluir que los modelos de accidentes no son fácilmente extrapolables a otras realidades o ciudades. A su vez, este tipo de modelos no son replicables en periodos de tiempo más específicos como fue el caso de los periodos punta mañana, tarde y medio día en donde los resultados de ajustes bajaron de manera importante y no se lograba una correcta estimación. La explicación de este resultado radica en la baja cantidad de accidentes en periodos acotados de tiempo así como también en que en periodos punta existen diferencias en los comportamientos de los patrones de causalidad, ya que el los conductores y las condiciones del tránsito son distintas al resto/promedio del día.

Cuando se comparan los resultados del periodo punta tarde con los periodos punta mañana y medio día se obtiene un resultado distinto, pues a pesar que este periodo presenta más accidentes que los otros dos, y a su vez, menos intersecciones de la base de datos con cero accidentes, los ajustes de los modelos replicados en el periodo punta tarde son menores que los dos otros periodos punta. Esto implica que en el periodo punta tarde es donde las características captadas a través de las variables definidas de operación e infraestructura explican en menor grado la ocurrencia de los accidentes. En otras palabras, los aspectos fuera del alcance de este estudio adquieren mayor importancia en este periodo, lo cual es motivo de una interesante línea de investigación futura.

Resumiendo, de las variables analizadas se observó que en los tres tipos de modelos estimados (lineales, Poisson y no lineales) aquellas que resultaron significativas y que generaron los mejores modelos fueron las variables de operación, específicamente el flujo total de la intersección, la densidad peatonal y el número de cruces peatonales, y en cuanto a las variables asociadas a la infraestructura, el número de pistas y el ángulo de la intersección resultaron como variables significativas. Ahora bien, la variable fundamental que explica la ocurrencia de accidentes es el flujo total de la intersección. Esto se concluye porque en todos los modelos ajustados la variable del flujo resultó significativa y con elevados test t, además de obtener los mejores ajustes de R^2 en los modelos donde se consideraba dicha variable.

Finalmente, y realizando las pruebas en modelos desagregando la variable de flujo vehicular total por ramas de la intersección, por movimientos y por tipología de vehículos, así como también utilizando variables de proporciones de flujo entre flujo principal y secundario, se observó que los ajustes de los modelos no mejoraban significativamente. Luego la utilización de la variable flujo total de la intersección resulta adecuado y más simple para la estimación de los modelos de accidentes en intersecciones urbanas.



CAPÍTULO 7 CONCLUSIONES

El objetivo principal de esta investigación fue identificar los factores que explican la ocurrencia y el número de accidentes en las intersecciones urbanas del Gran Concepción. Para ello se revisó el estado del arte, se definió un diseño experimental, luego se sintetizó una base de datos con distintas fuentes de información incluyendo la toma de datos en terreno, se analizaron estadísticamente las variables de modo tal de definir formalmente los patrones de comportamiento y las relaciones de éstas con la ocurrencia de accidentes y, finalmente se realizó una modelación matemática buscando replicar y estimar el número de accidentes en intersecciones urbanas.

Se observó que los modelos utilizados en la estimación de accidentes en intersecciones urbanas se clasifican en modelos Probit, regresiones y lineales generalizados. Del análisis de los modelos, el objetivo principal fue identificar las variables con las que se explicaba la ocurrencia de accidentes. De las que se identificaron, se realizó una clasificación en tres grupos; las variables de infraestructura, de operación y del entorno. Se concluyó que las relacionadas con el flujo vehicular son fundamentales para un modelo de estimación de accidentes, lo que se justifica, pues estaban presentes en todos los modelos revisados (significativas estadísticamente). No obstante, se encontraban en distintos formatos, es decir, en algunos casos, como flujo vehicular total, en otros, como flujos por ramas (principal y secundaria) y en ciertos modelos, como un cociente entre los flujos principal y secundario.

En el diseño experimental se definió el listado de variables iniciales de infraestructura (13), operación (19) y entorno (4) que se calcularían para la conformación de la base de datos. A su vez, la matriz factorial completa resultó ser de órdenes de magnitud muy elevados (10^{15} celdas). Esto determina una necesidad de simplificar el diseño factorial, descartando variables, disminuyendo en donde se pueda su cantidad de niveles e identificando todas las combinaciones no factibles entre atributos, con el objetivo de disminuir el tamaño de la matriz y poder realizar un modelo válido considerando la cantidad de datos que se tiene.

Para abordar el análisis se elaboró una base de datos que contiene 273 intersecciones, en donde para cada intersección se tienen los números de accidentes totales por año, y todos los valores de

las 36 variables independientes definidas. Esto se realizó también para los distintos periodos representativos con el objetivo de revisar diferencias entre periodos. En esta labor resultó importante la obtención de flujos vehiculares en las intersecciones, pues al ser la característica a priori más importante para estos modelos, este dato determinó en parte las intersecciones a analizar. En cuanto a los accidentes, se obtuvo una base de datos total de 3431 lo cual representa en promedio el 45% de los accidentes totales en intersecciones del Gran Concepción. Además, se obtuvo el número de accidentes por año para cada intersección de modo tal de en un futuro análisis poder observar el efecto del tiempo en los accidentes.

El proceso de análisis de las variables permitió descartar algunas para llegar al resultado final, en donde se obtuvieron: 15 variables de operación significativas, seis de infraestructura y solo una variable significativa asociada al entorno. Las variables de infraestructura que resultaron significativas son; el número de pistas por rama, la presencia de pista exclusiva de giro a izquierda en la rama principal, la presencia de mediana en la rama principal, el ángulo de la intersección y el número de calles unidireccionales. Las variables de operación significativas fueron casi todas las relacionadas con el flujo vehicular, a excepción del flujo opuesto y del flujo de taxi colectivos, pues estas variables presentaban muchos ceros en sus valores. Así mismo, el número de fases del semáforo, la densidad peatonal, el número de cruces peatonales y la visibilidad resultaron significativas estadísticamente, pero con un menor factor de correlación que las del flujo.

Si bien los modelos de regresiones tanto lineales como no lineales resultan con ajustes adecuados, estos alcanzan valores de R^2 (ajustado) de 0,45 y 0,52 respectivamente. Esta diferencia es acorde a lo esperado, pues por la naturaleza de los accidentes, un modelo no lineal resulta más ad-hoc para su estimación.

De este trabajo se obtuvo como resultado fundamental, saber qué variables tienen una mayor capacidad de explicación de la ocurrencia de accidentes. En este sentido las variables relacionadas con la operación, y principalmente las de flujo vehicular resultaron como las variables de mayor importancia en todos los modelos probados. En otras palabras, siempre fueron significativas estadísticamente, siempre se obtuvo elevados valores del test t y valores p muy pequeños (>0.01) asociados a éstas variables. Además de estas variables, también resultaron significativas variables

como el número de pistas, la visibilidad, el ángulo de la intersección, la densidad peatonal, el número de cruces peatonales y el número de paraderos. Estas últimas, si bien se repitieron en muchos modelos estimados, no presentaron el mismo peso relativo que aquellas del flujo. No obstante generaban una mejora en la capacidad predictiva del modelo y además resultaron significativas estadísticamente en varios modelos probados.

Finalmente el modelo estimado que resultó más adecuado fue el del tipo no lineal y la forma funcional de éste resultó ser similar al modelo de la AASHTO y a los modelos de Exposición al riesgo, es decir, modelos con una componente exponencial. En este caso, el modelo estimado obtuvo un ajuste de R^2 de 52% y además se simplificó la cantidad de variables a cinco (Flujo vehicular total de la intersección, Número de pistas de la rama secundaria, Angulo de la intersección, Densidad peatonal y Número de cruces peatonales).

Del análisis por periodo realizado se obtuvo que, el punta tarde es donde existe una mayor cantidad de accidentes por hora y una menor cantidad de variables significativas con relación a la ocurrencia de accidentes. Esto puede implicar que existen causas no consideradas en las variables medidas, las cuales adquieren mayor relevancia en este periodo de modelación.

A su vez, las variables que no resultaron significativas estadísticamente, podrían ser re-definidas, o modificadas esperando captar de mejor manera sus respectivos efectos sobre los accidentes. Por ejemplo, la variable “presencia de mediana”, podría cambiarse por una variable que considerase el ancho de las medianas, y/o “la presencia de pistas exclusivas de viraje”, podría convertirse en el número de pistas de viraje exclusivos en la intersección. En otras palabras, cuando del análisis de significancia se tiene un resultado negativo, no se debe descartar por completo tal variable, más bien, debe probarse re-definirla.

Se observó también que existe una deficiencia en la calidad datos de accidentes lo cual afecta la profundidad de los análisis que se puedan realizar. Esto pues no se tenía completa precisión sobre sus ubicaciones, severidades, tipos de accidentes e involucrados en éstos. Resulta entonces necesario para poder tener mejores análisis a futuro estandarizar la toma del dato de accidentes, según las siguientes recomendaciones:

- Ubicación exacta del accidente: Además de definir calles de la intersección, es importante saber si ocurrió justo en ésta o en alguna de las calles mencionadas y aguas abajo o arriba de la intersección misma.
- Tipo de accidente: Se refiere a estandarizar y homologar la nomenclatura a una en donde se pueda conocer los actores involucrados (veh-veh, veh-peatón, veh-no motorizado).
- Severidad: Existen inconsistencias en los niveles intermedios. Luego, es necesario capacitar a quienes toman el dato y, además, re-definir las categorías con el nombre oficial de la escala KABCO, los cuales resultan más explicativos e intuitivos que, por ejemplo, un accidente “medianamente grave”, evitando así la subjetividad en la toma del dato.

Además, como resultado de esta investigación, la obtención de datos precisos debe estar centradas en aquellas variables que resultaron ser las de mayor importancia para la predicción de la cantidad de accidentes.

Finalmente, existen líneas de investigación en las cuales ahondar. Por ejemplo, determinar cómo interactúan entre sí las distintas variables independientes con respecto al número de accidentes. Otro aspecto interesante sería revisar la significancia de las variables en las distintas severidades de accidentes, de tal manera de estudiar cómo cambian la importancia de éstas en sus efectos sobre la ocurrencia de accidentes de severidades distintas. Además, si se complementara la base de datos generada con información de intersecciones no semaforizadas, se podrían analizar otro tipo de variables, como “el tipo de control” y así ver cómo afecta al número de accidentes. Del mismo modo, se podría esperar que variables como las pistas de virajes exclusivas, que en este análisis no resultaron significativas, en uno análogo, pero con intersecciones no semaforizadas, resultaran de importancia en el efecto sobre los accidentes.

Asimismo, futuras investigaciones debiesen ahondar en la modelación de los accidentes utilizando modelos más elaborado, así como también, incluyendo la modelación de accidentes como una probabilidad, generando así modelos probabilísticos y no determinísticos como los estimados en este primer trabajo. En esta línea, por ejemplo, a través de modelos de elecciones discretas, pero utilizando distribuciones de probabilidades “continuas” (Probit), lo que representa todo un desafío, hoy en día, así como un interesante enfoque para abordar en la modelación de accidentes, pues en

la literatura actual no existe un método definido o consensuado para calcular la componente de probabilidad en modelos de estimación de accidentes.



REFERENCIAS

Abdel, M. y J. Keller (2004) Exploring the overall and specific crash severity levels at signalized intersections. **Accident Analysis and Prevention**. **37**. 417-425.

American Association of State Highway and Transportation Officials (2010) **Highway Safety Manual**. Primera edición. Washington DC.

Anselin, L. y D., Arribas-Bel (2013) Spatial fixed effects and spatial dependence in a single cross-section. **Regional Science**. **92**. 3-17.

Castro, M., R. Paleti y C., Baht (2012) A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. **Transportation Research Part B**. **46**. 253-272.

Chen, C. y Y. Xie (2016) Modeling the effects of AADT on predicting multiple-vehicle crashes at urban and suburban signalized intersections. **Accident Analysis and Prevention**. **91**. 72-83.

Chias, L. (2004) **Atlas de la seguridad vial en México**. Instituto de Geografía, UNAM, Secretaría de Salud, Consejo Nacional para la Prevención de accidentes. Ciudad de México.

Cohen, J. (1988) **Statistical Power Analysis for the Behavioral Sciences**. 2a Edición. Academic Press. Nueva York.

Cunto, F., L. Sobreira y S. Ferreira (2014) Assessing the transferability of the Highway Safety Manual Predictive Method for Urban Roads in Fortaleza City, Brazil. **Journal of Transportation Engineering**. **2014(141)**. 04014072-1 - 04014072-8.

Fridstrom, L., J. Ifver, J. Ingebrigtsen y S. Kulmala, (1995) Measuring the contribution of randomness, exposure, weather and daylight to the variation in road accident counts. **Accident Analysis and Prevention**. **27**. 1-20.

GESITRAN (2014) **Plataforma tecnológica para la gestión integral del sistema de transporte de la Región de Biobío.** <http://www.gesitranbiobio.cl/>. Visitado el 20/12/2016.

Gold, P.A. (1998) **Using engineering to reduce traffic accidents.** Traffic Safety. The Inter American Development Bank. Estados Unidos.

Greibe, P. (2003) Accident prediction models for urban road. **Accident Analysis and Prevention.** **35.** 273-285.

Hartigan, J. A.; Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. **Journal of the Royal Statistical Society.** **28(1).** 100–108.

IBM (2014) **SPSS Statistics Base 22.0.** Software estadístico. Nueva York.

Kumara, S. S. P. y H.C. Chin (2003) Modeling Accident Occurrence at Signalized Tee Intersections with Special Emphasis on Excess Zeros. **Traffic Injury Prevention.** **4(1).** 53-57.

MathWave Technologies (2010) **EasyFit 5.5.** Software estadístico.

Minitab Inc. (2005) **MINITAB Release 14.** Software estadístico.

MINVU (2009) **Manual de vialidad urbana, recomendaciones para el diseño de elementos de infraestructura vial urbana.** Ministerio de Vivienda y Urbanismo. Chile.

Mitra, S. y S. Washington (2007) On the nature of over-dispersion in motor vehicle crash prediction models. **Accident Analysis and Prevention.** **39.** 459-468.

MOP (2015) **Manual de Carreteras Volumen N°3 Instrucciones y criterios de diseño.** Ministerio de Obras Públicas. Chile.

Navarro, A., F. Utzet, P. Puig, J. Caminal y M. Martin (2001) Negativa binomial distribution versus Poisson in the analysis of recurrent phenomena. **Gac Sanit.** **15 (5)**. 447-452.

Salifu, M. (2003) Accident prediction models for unsignalised urban junctions in Ghana. **IATSS Research.** **28**. 68-81.

Sawalha, Z. y T. Sayed (2006) Transferability of accident prediction models. **Safety Science.** **44**. 209-219.

SECTRA (2014) **Actualización de programaciones de semáforos de la región del Bío Bío**. Subsecretaría de Transporte. Concepción. Realizado por Aristo Consultores.

SECTRA (2016) **Estudio mejoramiento eje Camilo Henríquez y par Bulnes-Cruz, Concepción**. Subsecretaría de Transporte. Concepción. Realizado por Cipres Ltda.

Statpoints Technologies Inc. (2014) **Statgraphics Centurion XVII**. Software estadístico.

Tasic, I. y R. J. Porter (2016) Modeling spatial relationships between multimodal transportation infrastructure and traffic safety outcomes in urban environments. **Safety Science.** **82(2016)**. 325-337.

Vieira Gomes, S. (2013) The influence of the infrastructure characteristics in urban road accident occurrence. **Accident Analysis and Prevention.** **60**. 289-297.

Vieira Gomes, S., S., Reddy y D., Lord (2012) Estimating the safety performance of urban intersections in Lisbon, Portugal. **Safety Science.** **50**. 1732-1739.

Wang, C., M., Quddus y S., Ison (2011) Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model. **Accident Analysis and Prevention.** **43**. 1979-1990.

ANEXO 4.1: Periodización del Gran Concepción

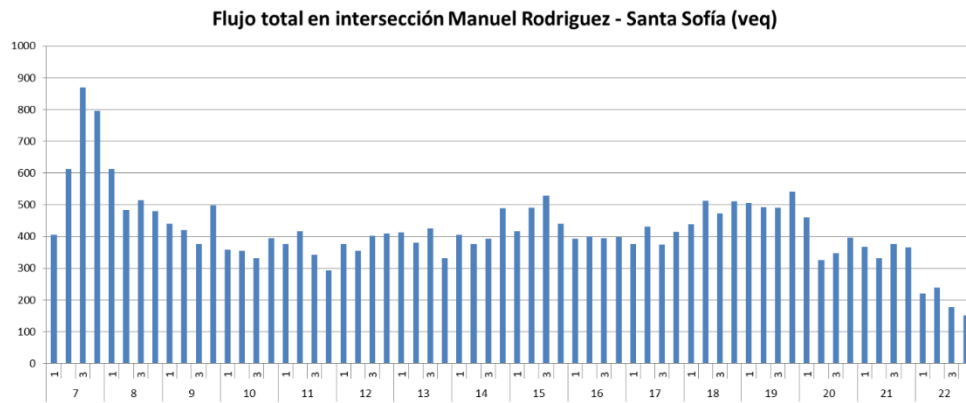


Figura A.4.1 Flujo total durante el día laboral en M. Rodríguez con Sta. Sofía, Chiguayante

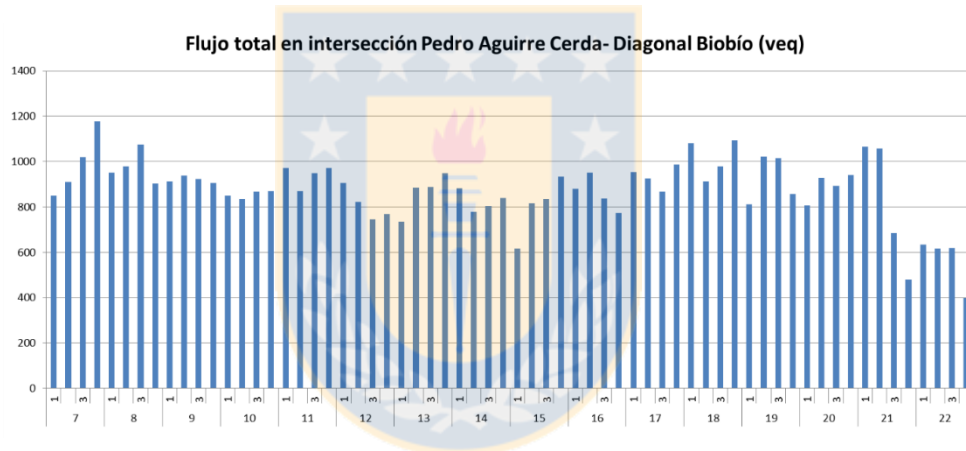


Figura A.4.2 Flujo total durante el día laboral en P. Aguirre Cerda con Diag. Biobío, San Pedro

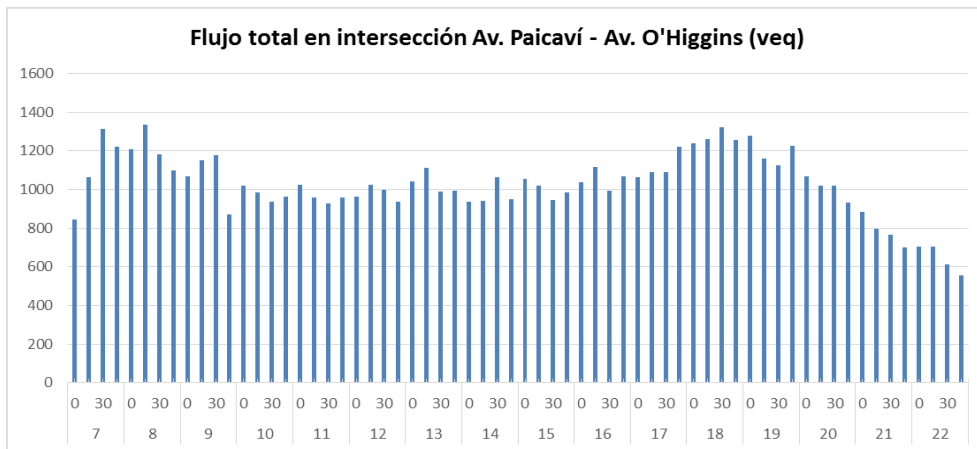


Figura A.4.3 Flujo total durante el día laboral en Av. Paicaví con Av. O'Higgins, Concepción

De los histogramas presentados y utilizando el método de la hora móvil más cargada se escogieron los siguientes periodos punta del día laboral.

Tabla A.4.1 Resultados de la periodización realizada, definiendo 2 horas en cada periodo

Intersección	Pedro Aguirre Cerda - Diagonal Biobío	Av. Paicaví - Av. O'Higgins	Manuel Rodríguez - Santa Sofía
PM	7:30-8:30	7:45-8:45	7:15-8:15
PMD	13:15-14:15	13:00-14:00	12:45-13:45
PT	18:00-19:00	18:15-19:15	18:45-19:45



ANEXO 4.2 Distribuciones espaciales de los accidentes de la muestra

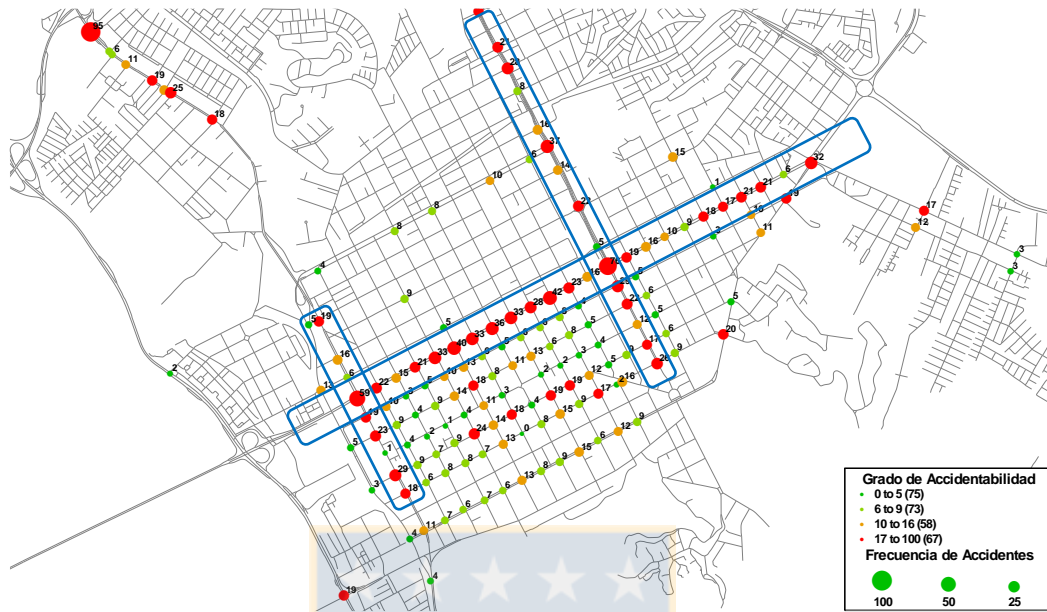


Figura A.4.4 Intersecciones y accidentes en la comuna de Concepción

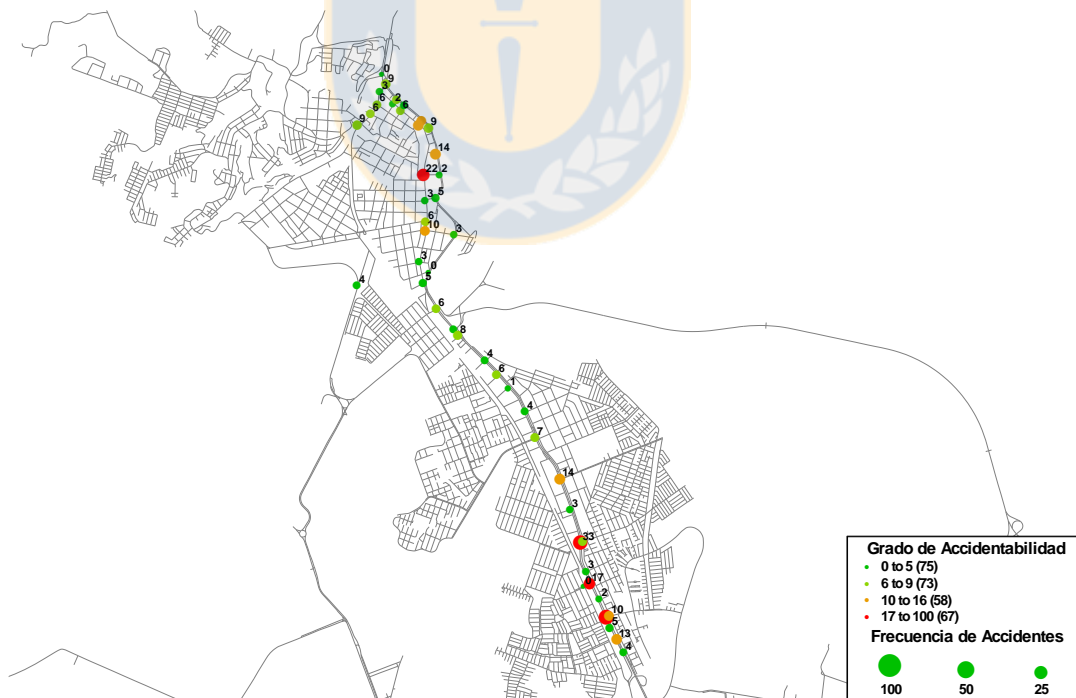


Figura A.4.5 Intersecciones y accidentes en la comuna de Talcahuano



Figura A.4.6 Intersecciones y accidentes en la comuna de San Pedro de la Paz

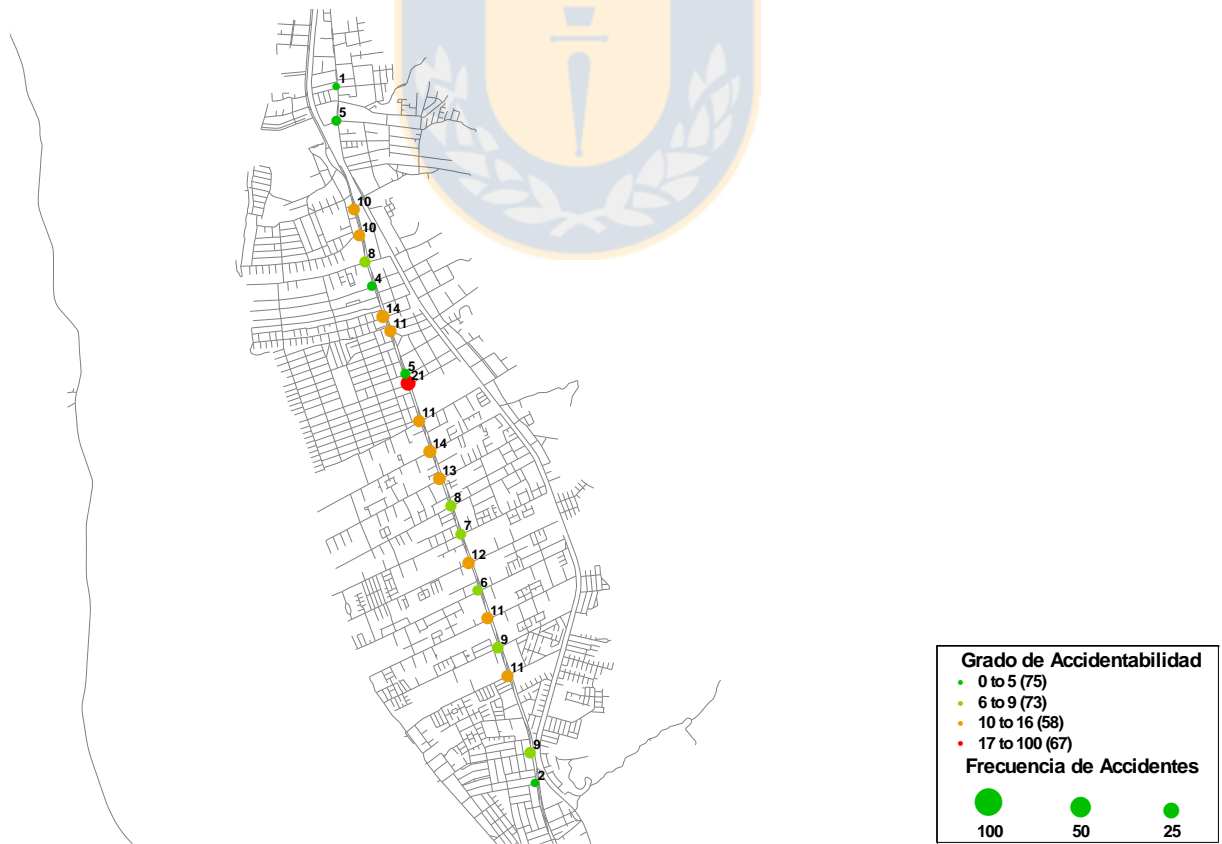


Figura A.4.7 Intersecciones y accidentes en la comuna de Chiguayante

ANEXO 5.1 Histogramas de frecuencias

Histogramas de las variables relacionadas al flujo por tipo de movimientos.

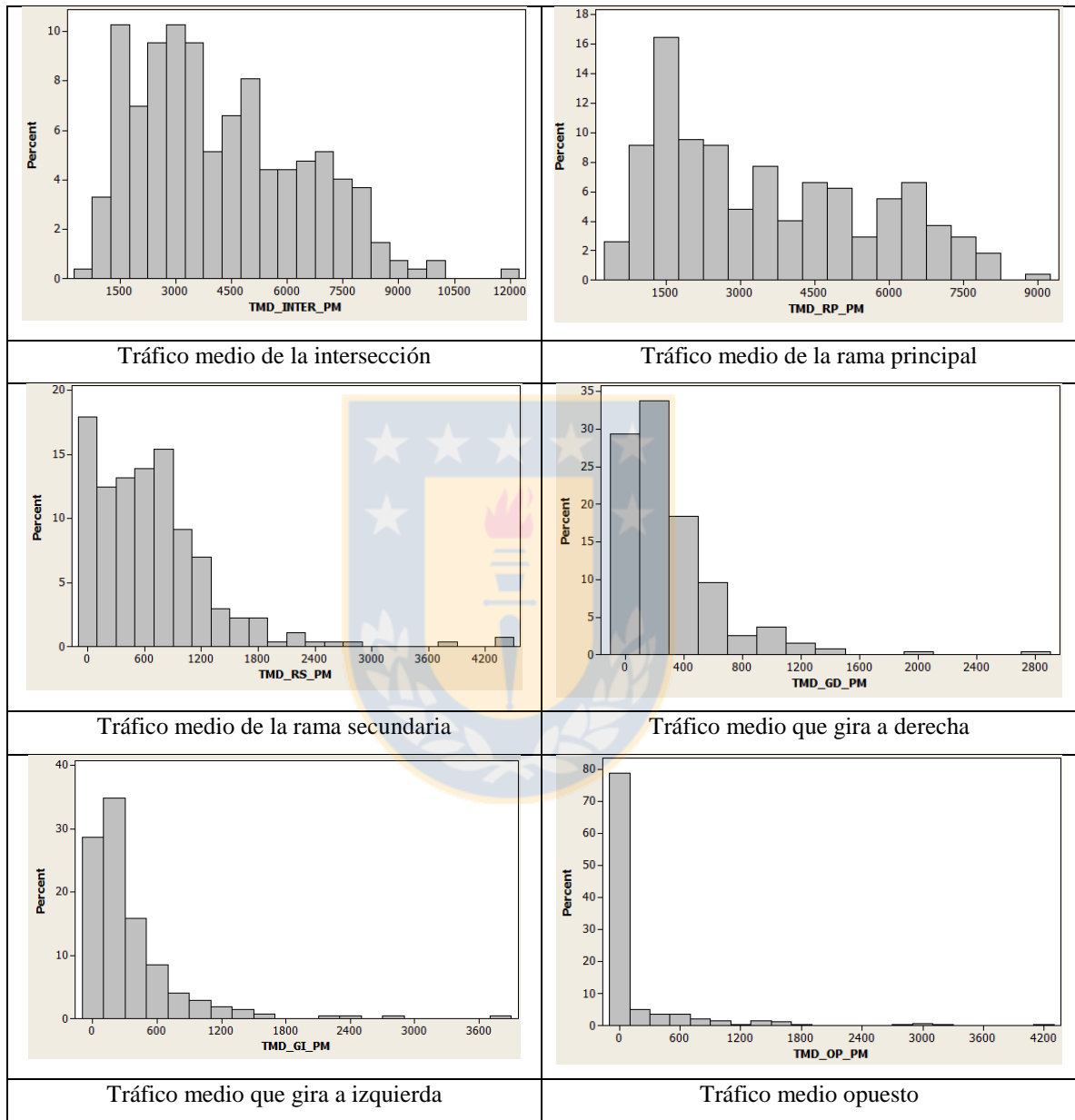


Figura A.5.1 Distribución de frecuencia de los flujos por tipo de movimiento del periodo PM

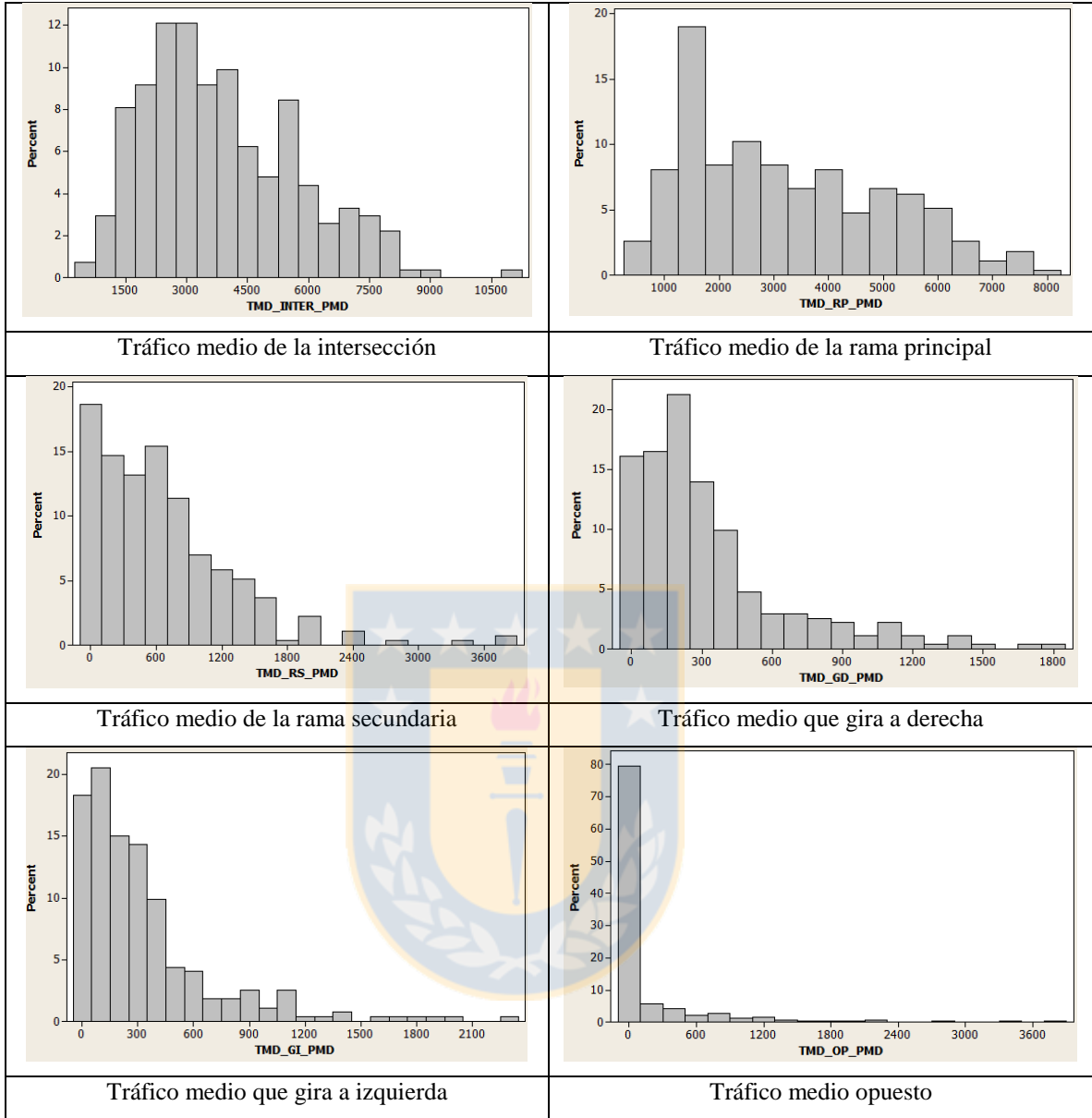


Figura A.5.2 Distribución de frecuencia de los flujos por tipo de movimiento del periodo PMD

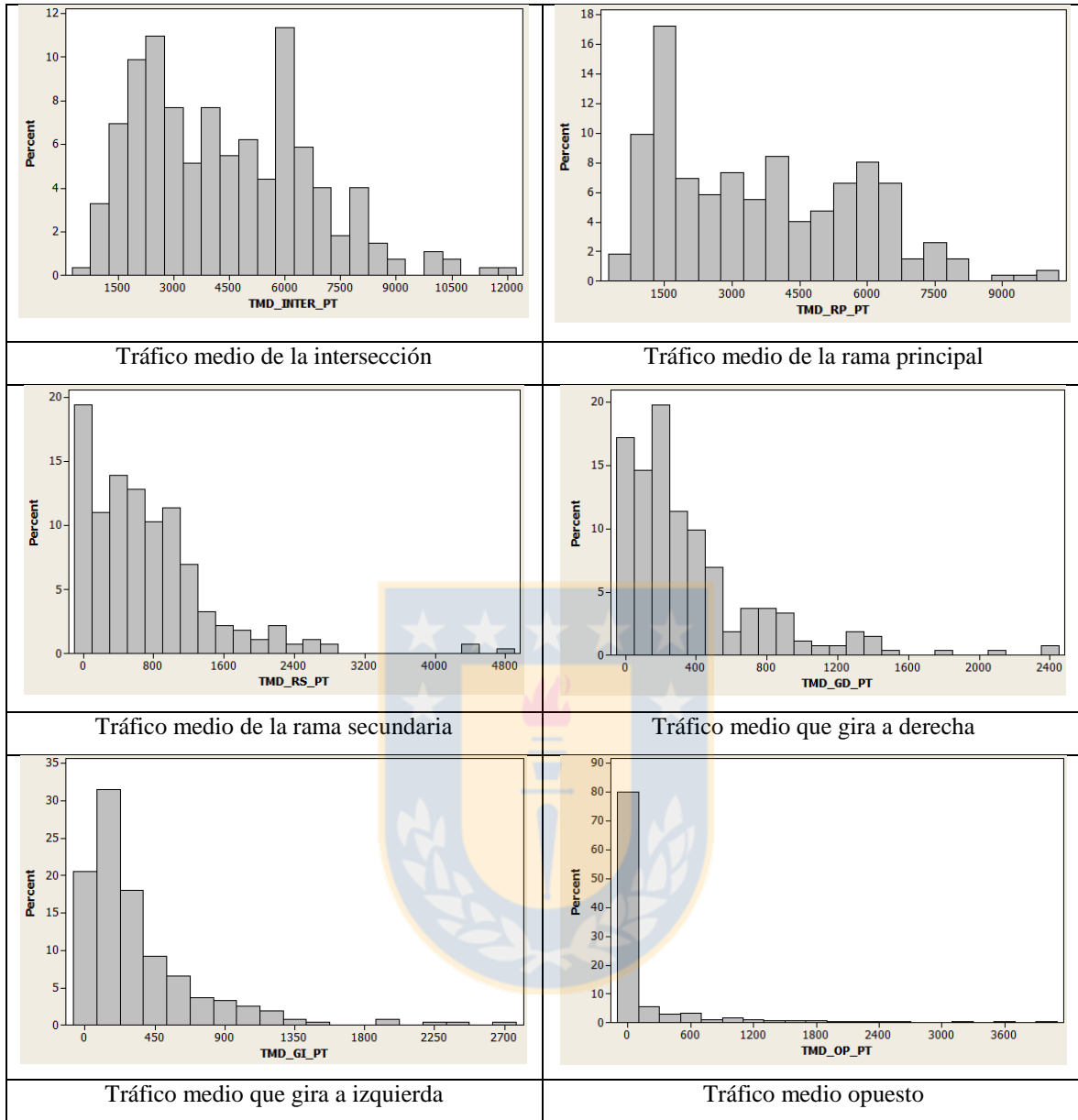


Figura A.5.3 Distribución de frecuencia de los flujos por tipo de movimiento del periodo PT

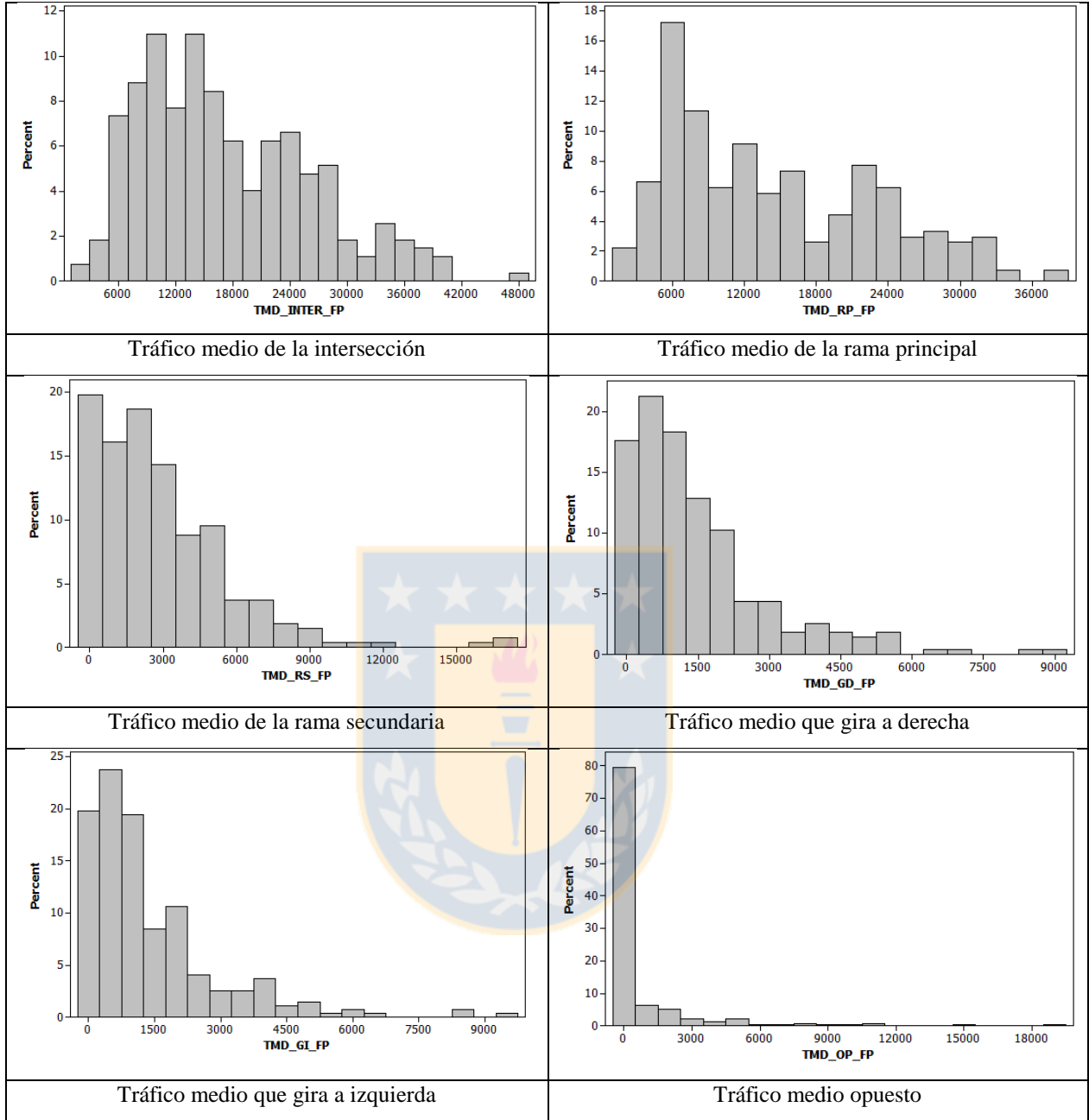


Figura A.5.4 Distribución de frecuencia de los flujos por tipo de movimiento del periodo FP

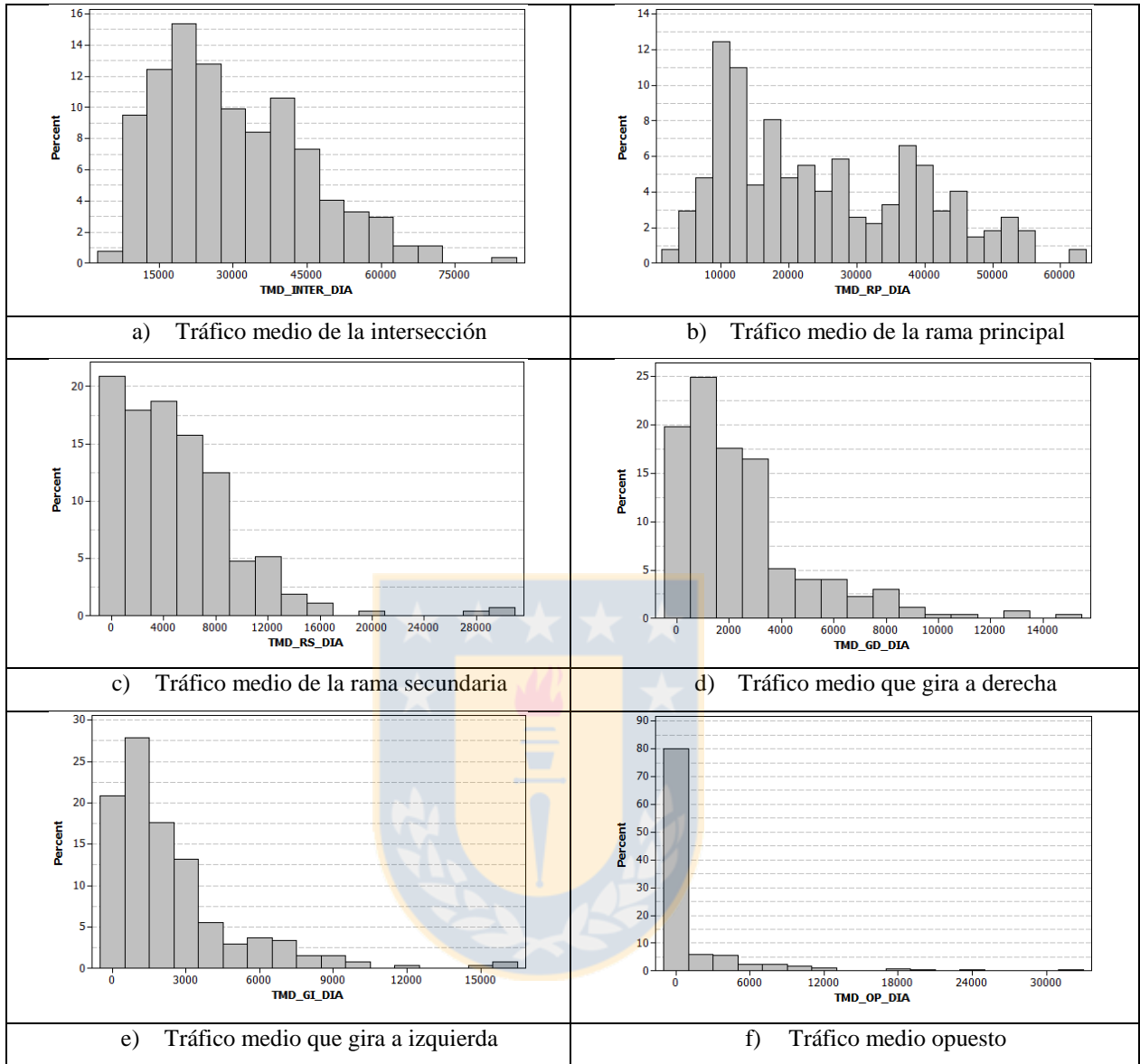


Figura A.5.5 Distribución de frecuencia de los flujos por tipo de movimiento del periodo diurno

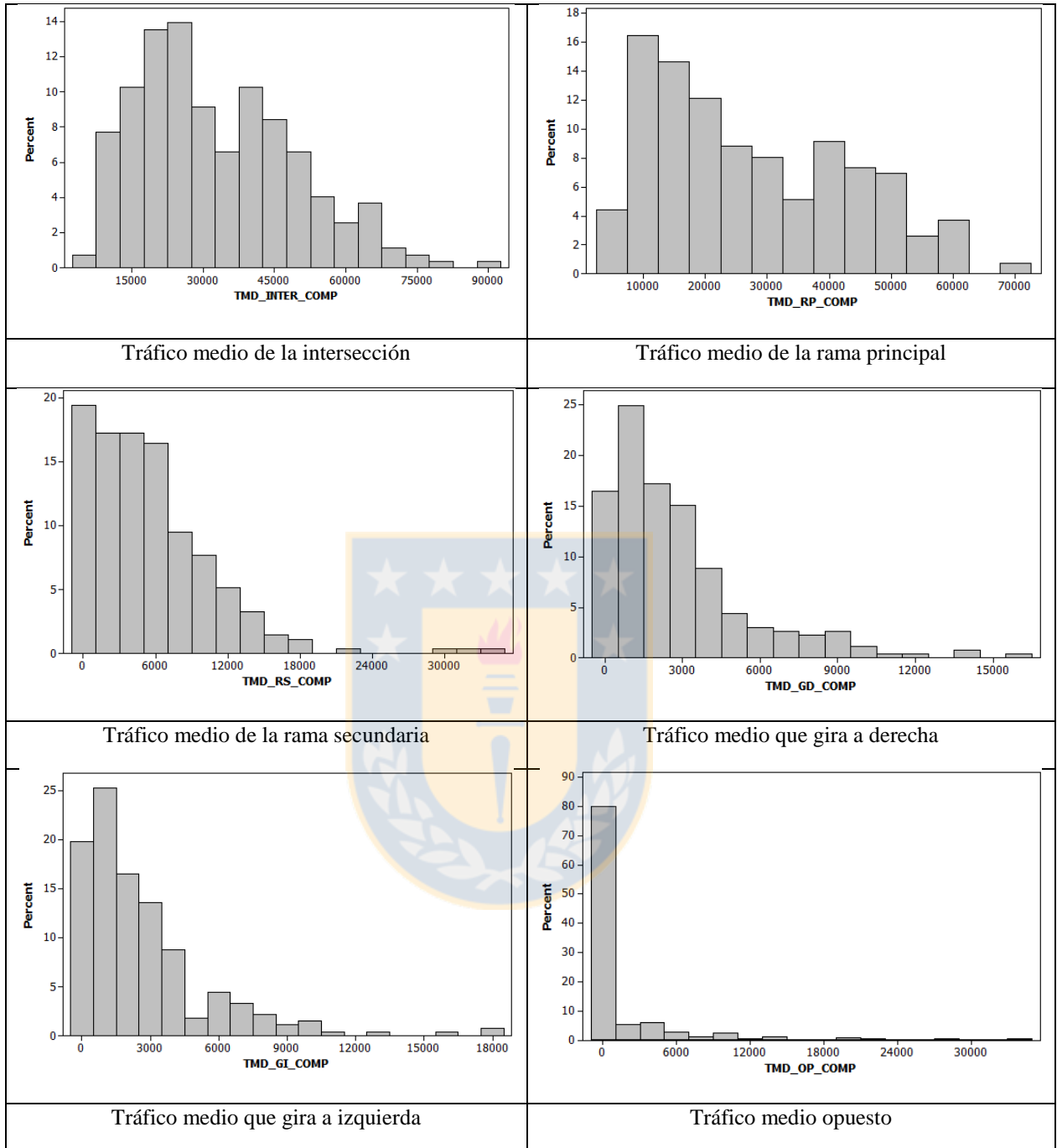


Figura A.5.6 Distribución de frecuencia de los flujos por tipo de movimiento del día completo

Histogramas de las variables relacionadas al flujo por tipo de vehículos.

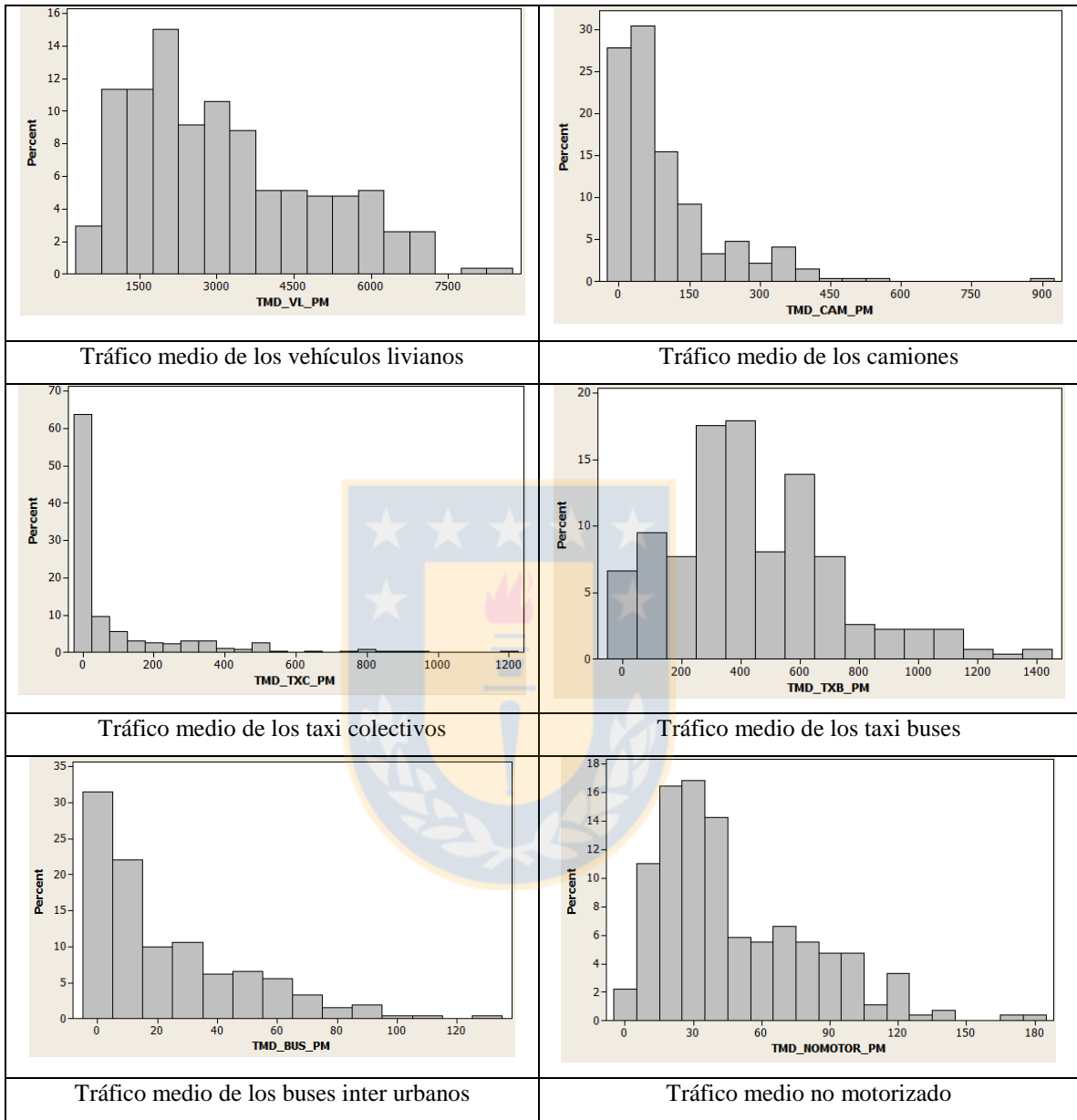


Figura A.5.7 Distribución de frecuencia de los flujos por tipo de vehículo del periodo PM

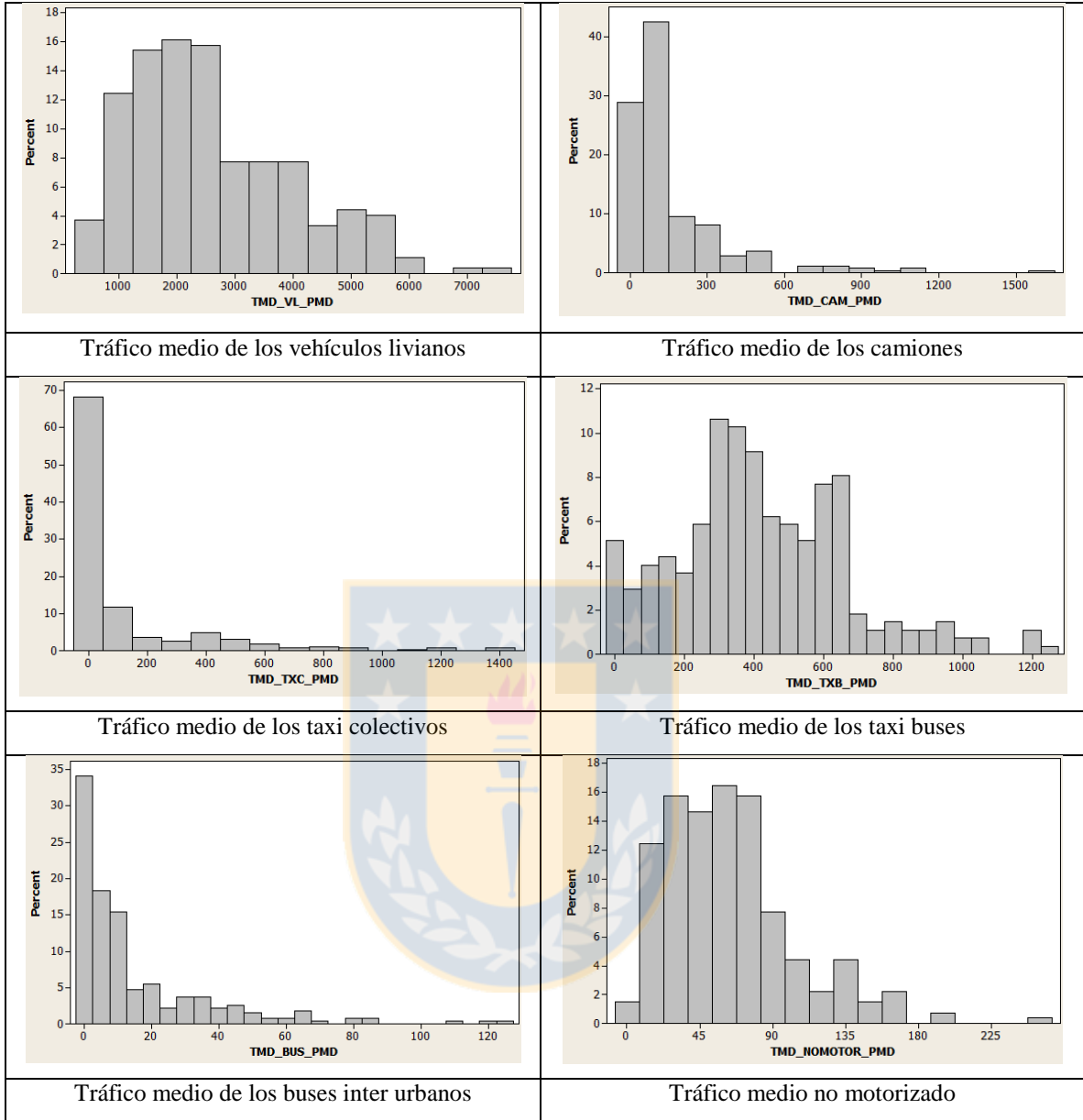


Figura A.5.8 Distribución de frecuencia de los flujos por tipo de vehículo del periodo PMD

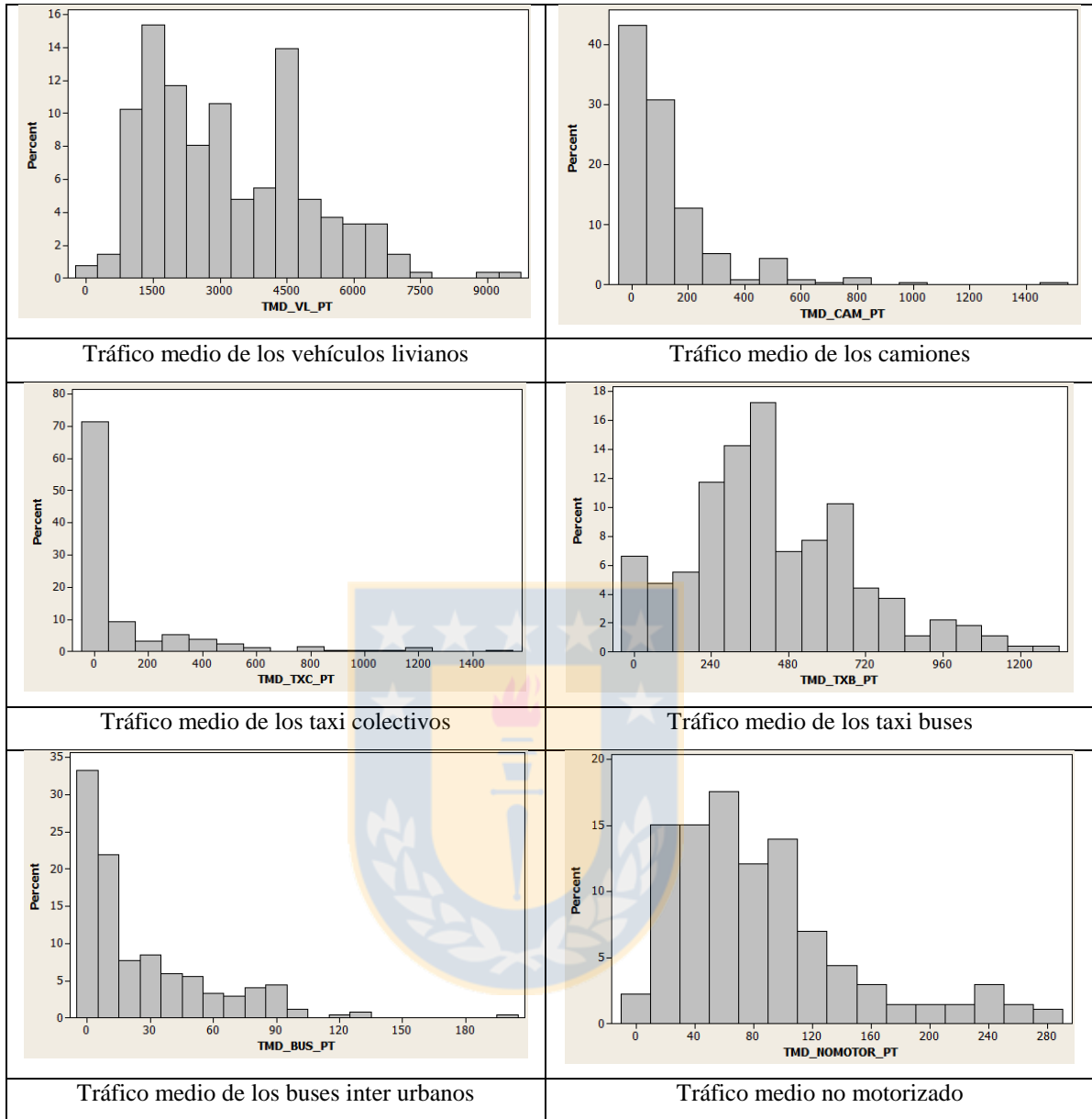


Figura A.5.9 Distribución de frecuencia de los flujos por tipo de vehículo del periodo PT

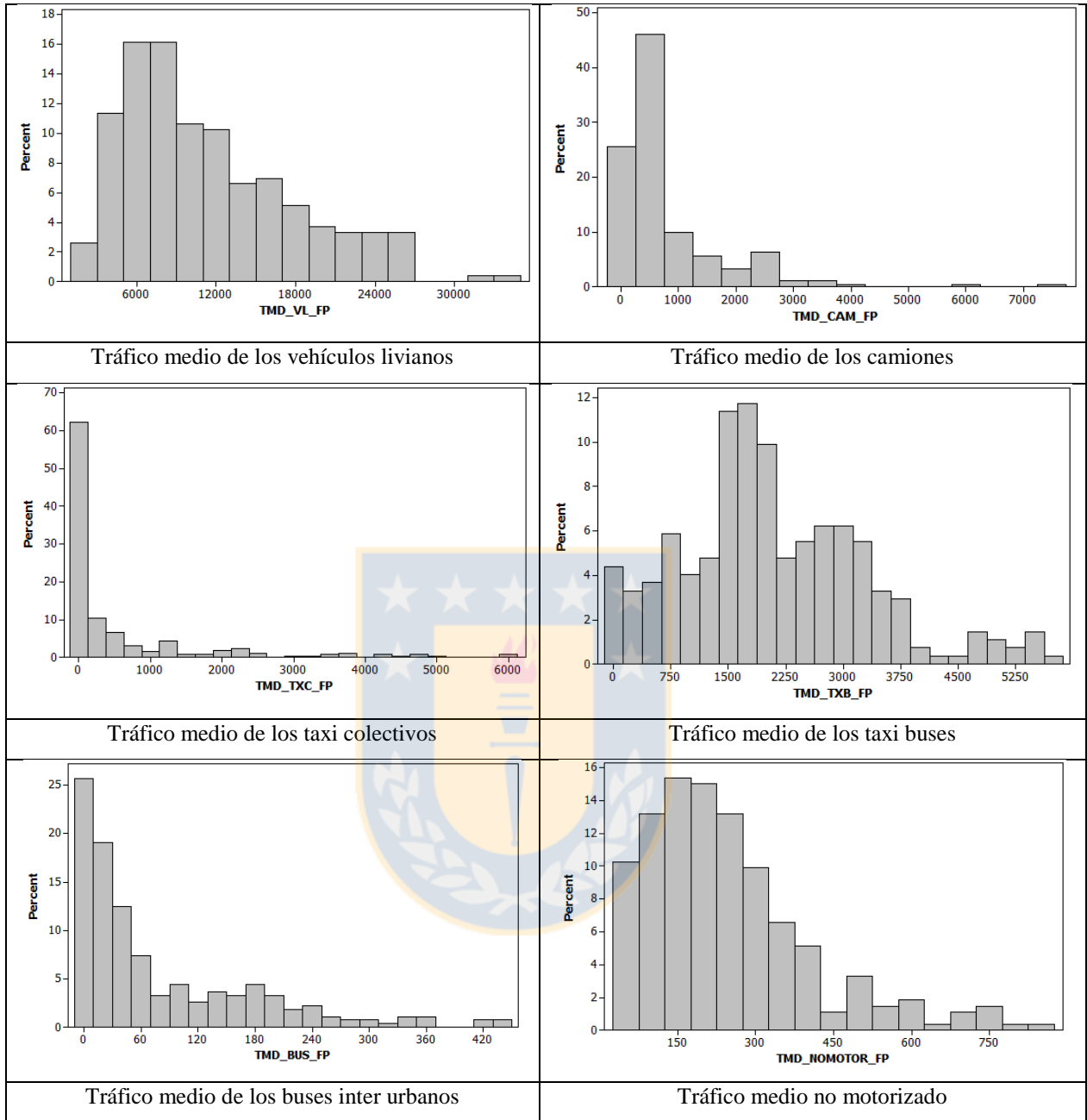


Figura A.5.10 Distribución de frecuencia de los flujos por tipo de vehículo del periodo FP

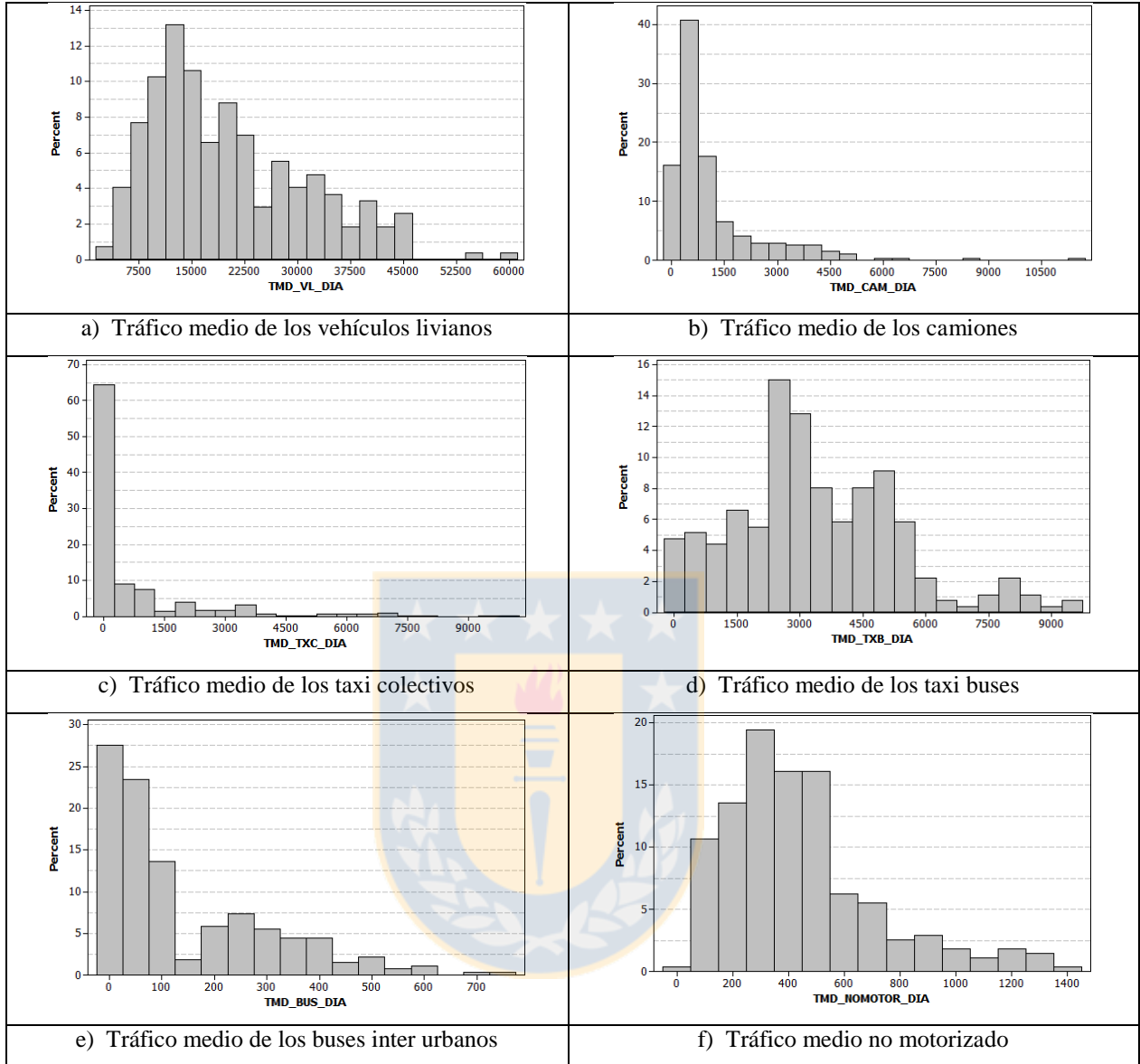


Figura A.5.11 Distribución de frecuencia de los flujos por tipo de vehículo del periodo diurno

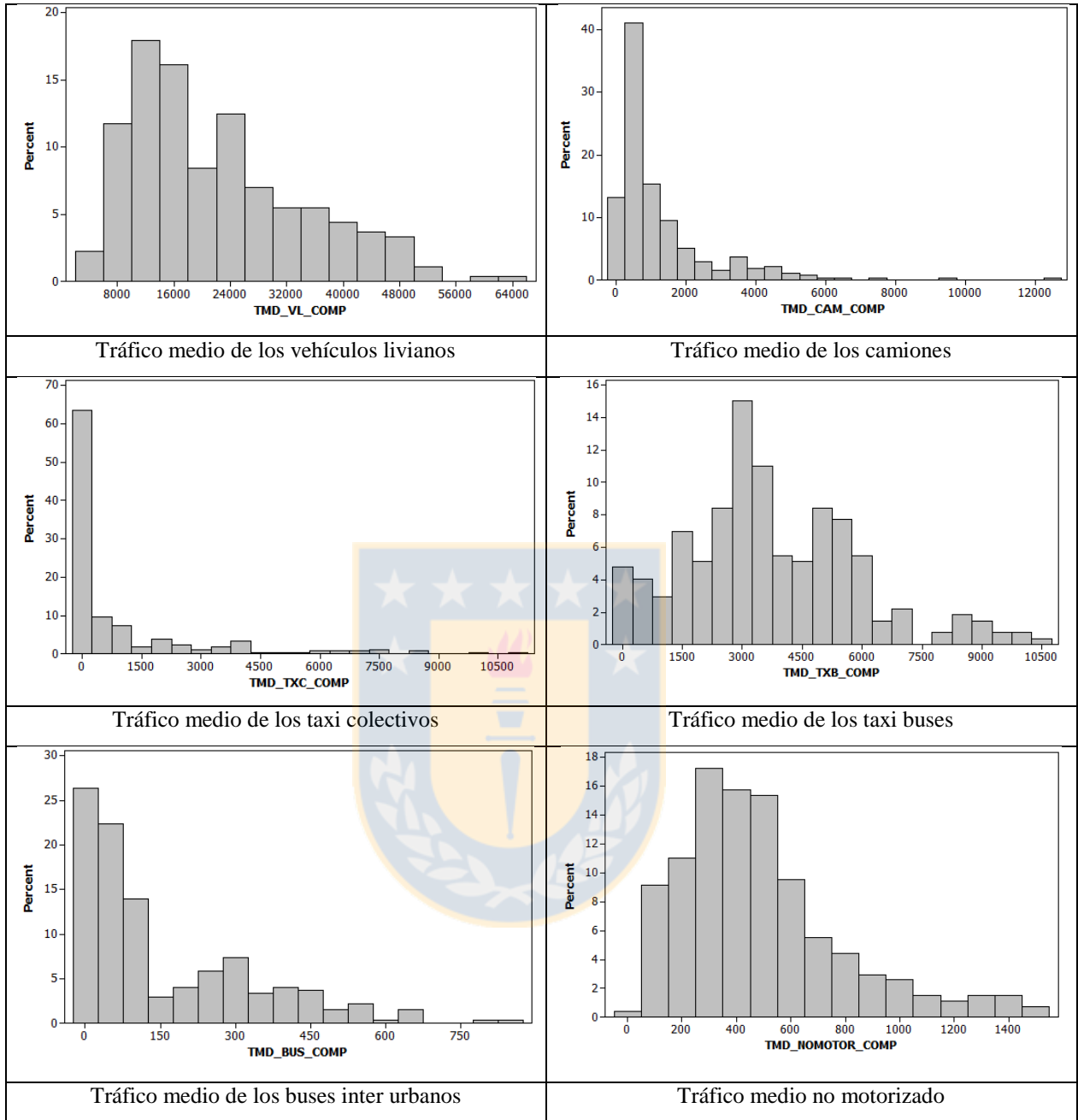


Figura A.5.12 Distribución de frecuencia de los flujos por tipo de vehículo del día completo

Histogramas de variables relacionadas a los cuocientes entre flujos vehiculares

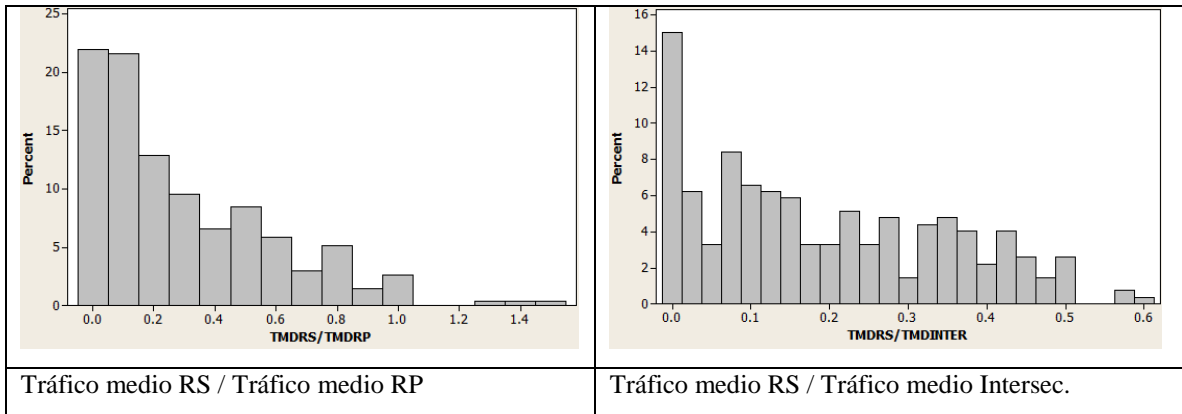


Figura A.5.13 Distribución de frecuencia de los cuocientes de flujos del periodo PM

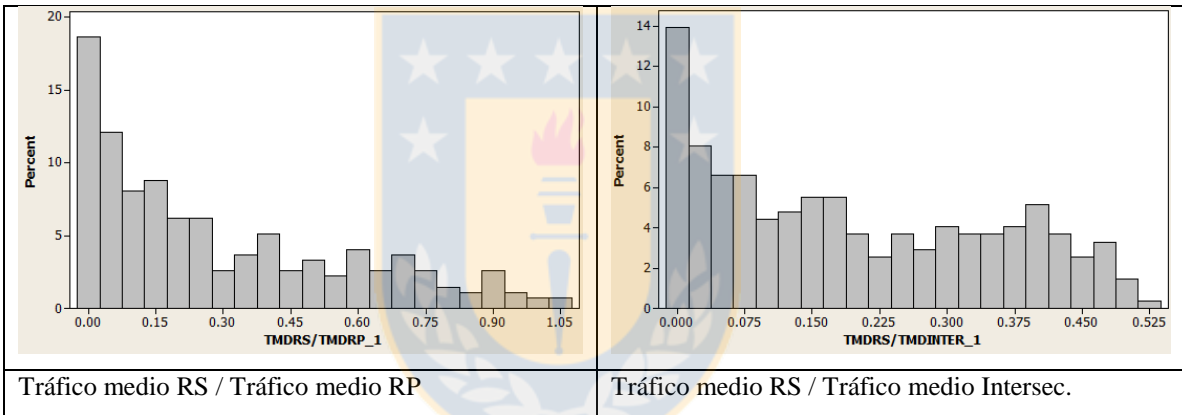


Figura A.5.14 Distribución de frecuencia de los cuocientes de flujos del periodo PMD

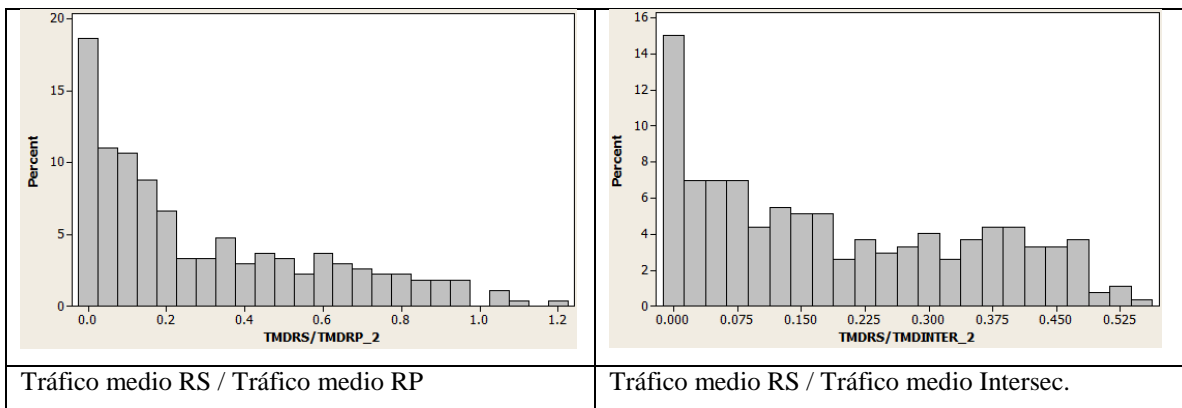


Figura A.5.15 Distribución de frecuencia de los cuocientes de flujos del periodo PT

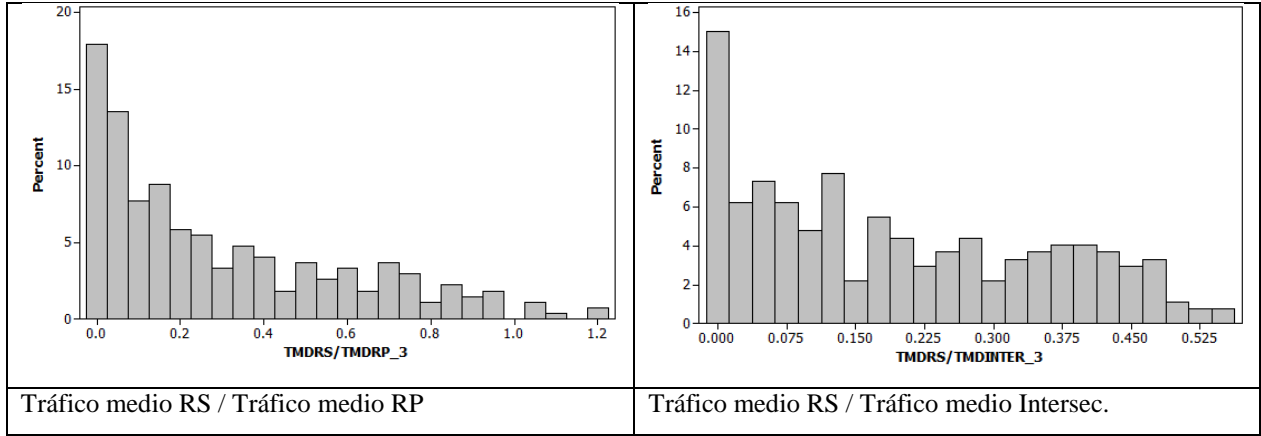


Figura A.5.16 Distribución de frecuencia de los cuocientes de flujos del periodo FP

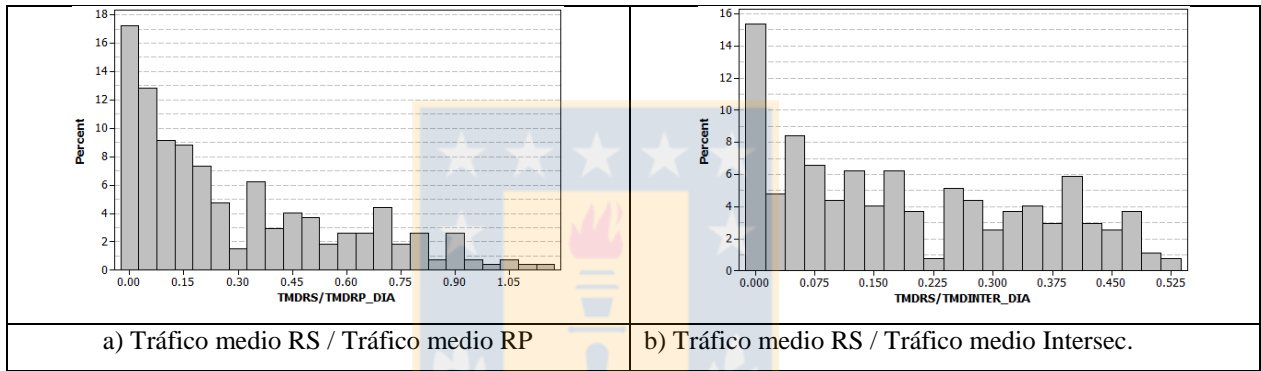


Figura A.5.17 Distribución de frecuencia de los cuocientes de flujos del periodo diurno

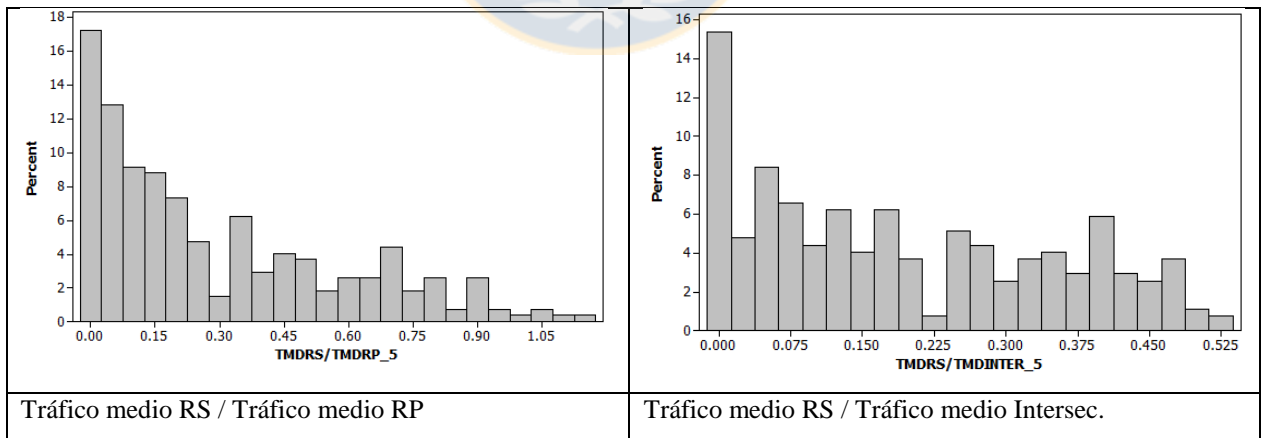


Figura A.5.18 Distribución de frecuencia de los cuocientes de flujos del día completo

ANEXO 5.2 Efectos sobre la media de accidentes de las variables analizadas

Se presenta en este anexo, los resultados gráficos del análisis de los efectos principales sobre la media de accidentes de cada grupo definido. Para ello se presenta a continuación los resultados y un pequeño análisis para cada grupo del periodo punta mañana. Los demás periodos también son reportados y presentados todos los respectivos gráficos.

Se observa que del grupo de variables relacionadas al flujo por movimientos, el tráfico de la intersección, el tráfico que gira a izquierda y derecha y el tráfico en la rama principal tienen un comportamiento creciente continuo. En cambio, la variable de tráfico opuesto no presenta un efecto significativo. Por otro lado, la variable de tráfico en la rama secundaria tiene un comportamiento anómalo sobre la gran media, pues en el último tramo presenta un aumento elevado con respecto a los demás nivel donde no hay un cambio significativo en la media.

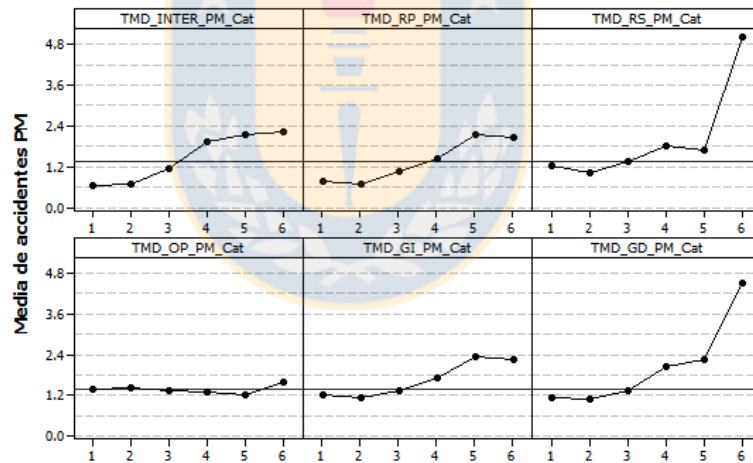


Figura A.5.19 Efectos del flujo por movimientos sobre la media de accidentes en PM

Se observa que las variables asociadas al tráfico medio de vehículos livianos, de camiones y de taxi buses resultan significativas sobre la media de accidentes, aumentando el número de estos a medida que aumenta el flujo. En cambio, las variables de tráfico medio taxi colectivos, de buses interurbanos y de tráfico no motorizado no generan un cambio que resulte significativo sobre la media de accidentes.

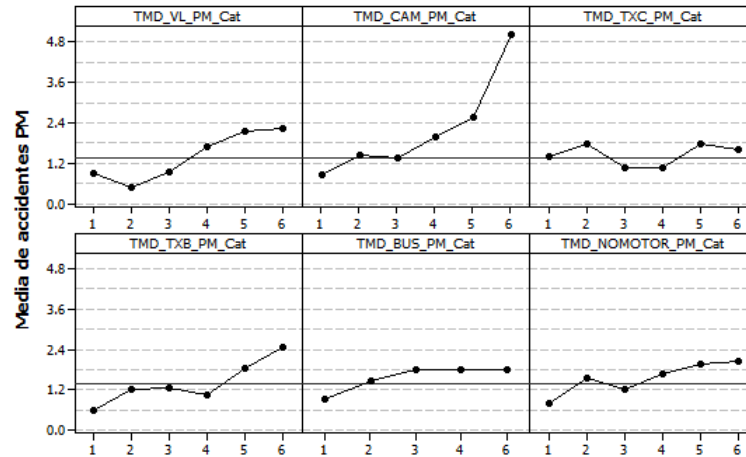


Figura A.5.20 Efectos del flujo por tipología sobre la media de accidentes en PM

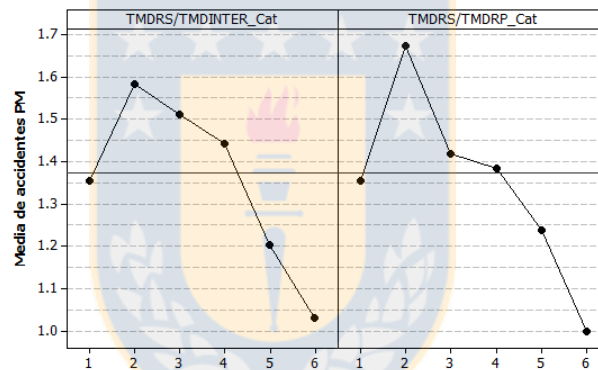


Figura A.5.21 Efectos de los cuocientes del flujo sobre la media de accidentes en PM

Se observa que de las variables analizadas, sólo el número de cruces peatonales, el número de cruces peatonales con refugio y el número de paraderos resultan significativos en sus efectos sobre la gran media de accidentes.

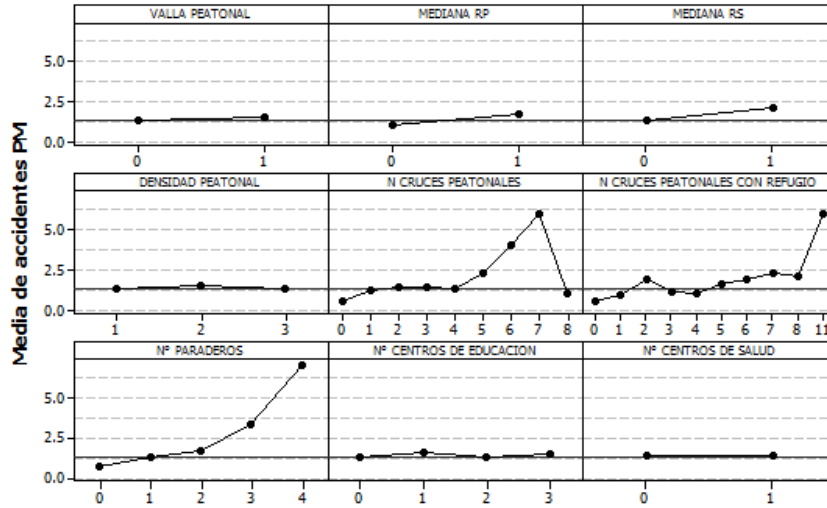


Figura A.5.22 Efectos de las variables peatonales sobre la media de accidentes en PM

Se observa que la presencia de giros a izquierda por sí sola no resultan significativas en su efecto sobre la media de accidentes, no obstante las presencias de pistas exclusivas de giros tanto a derecha como a izquierda resultan significativas y en intersecciones donde existen pistas exclusivas, existe una mayor cantidad de accidentes.

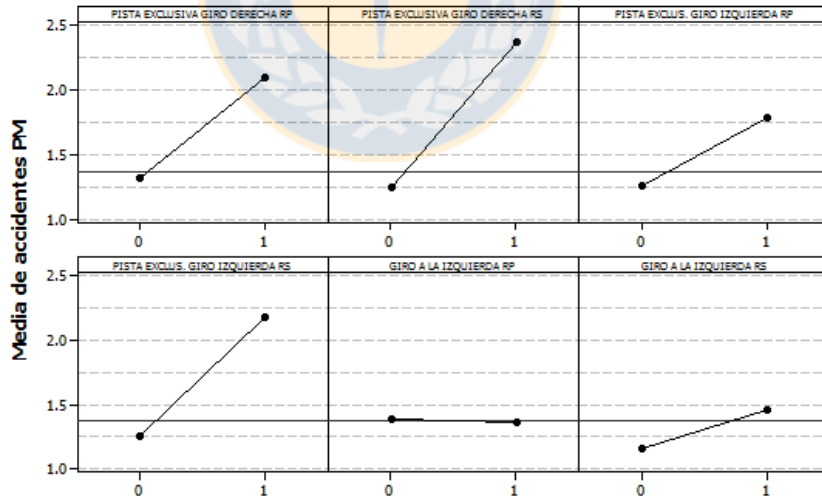


Figura A.5.23 Efectos de los giros sobre la media de accidentes en PM

Se observa que la variable número de pistas de la rama secundaria resulta significativa en cuanto a su efecto sobre la media de accidentes, aumentando el número de éstos a medida que aumenta el número de pistas. La variable número de calles unidireccionales presenta tendencia a disminuir la

cantidad de accidentes a medida que más calles son unidireccionales, pero su efecto es marginal. Finalmente, las variables número de ramas de acceso y número de pistas en rama principal no resultan significativas.

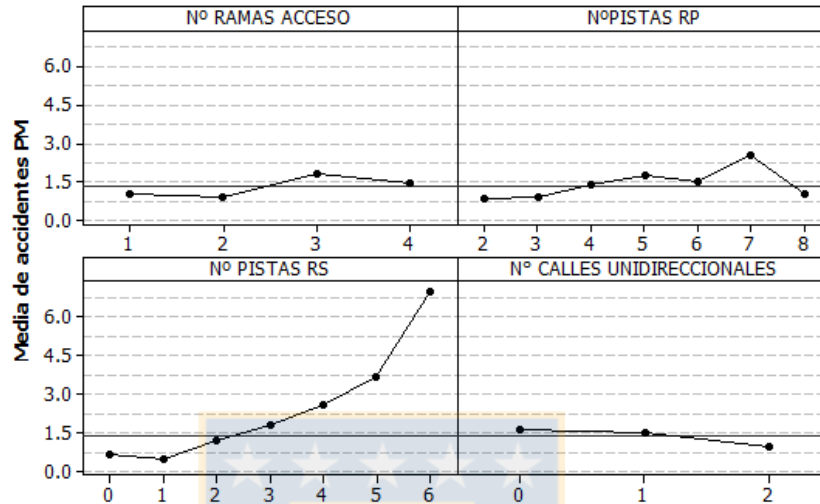


Figura A.5.24 Efectos del diseño geométrico sobre la media de accidentes en PM

Se observa que existen en los datos una gran proporción de intersecciones con 90° . No obstante, se observa que los accidentes en intersecciones con ángulos de intersección entre 0 y 30° resultan reducidos, no obstante se observa un punto extremo en el ángulo de 40° con un elevado número de accidentes. De esta variable no se puede concluir que sea significativa pues su comportamiento no presenta una tendencia clara. La variable de visibilidad presenta significancia sobre la media, siendo el nivel 4 (visibilidad entre 68.5 m y 95.3 m) el rango en donde mayor cantidad de accidentes existe. Finalmente la variable visión de intersección (categórica) no presenta significancia sobre la media de accidentes.

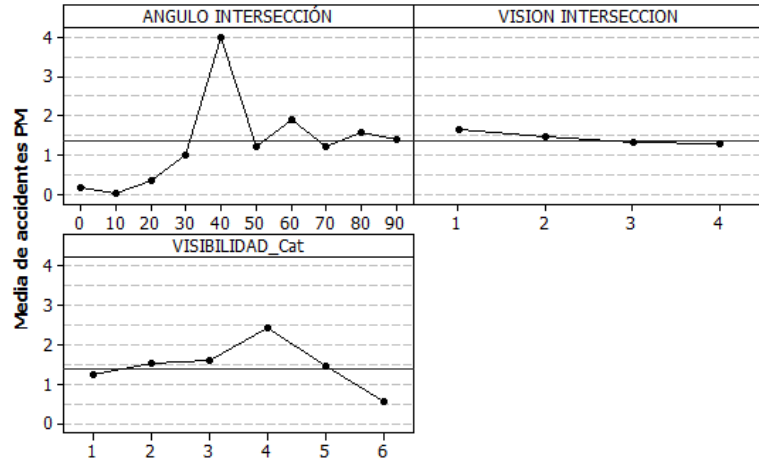


Figura A.5.25 Efectos de la visibilidad sobre la media de accidentes, base de datos PM

Se observa que la variable tipo de intersección tiene significancia sobre la media de accidentes, disminuyendo el número de accidentes en intersecciones del tipo peatonales (nivel 3) y siendo mayor en aquellas del tipo T (nivel 1). En lo que respecta al número de fases del semáforo se observa que cuando son tres fases la accidentabilidad aumenta. Finalmente, el tipo de zona también resulta significativo sobre la media de accidentes, siendo la zona comercial (nivel 2) en donde existen menos accidentes.

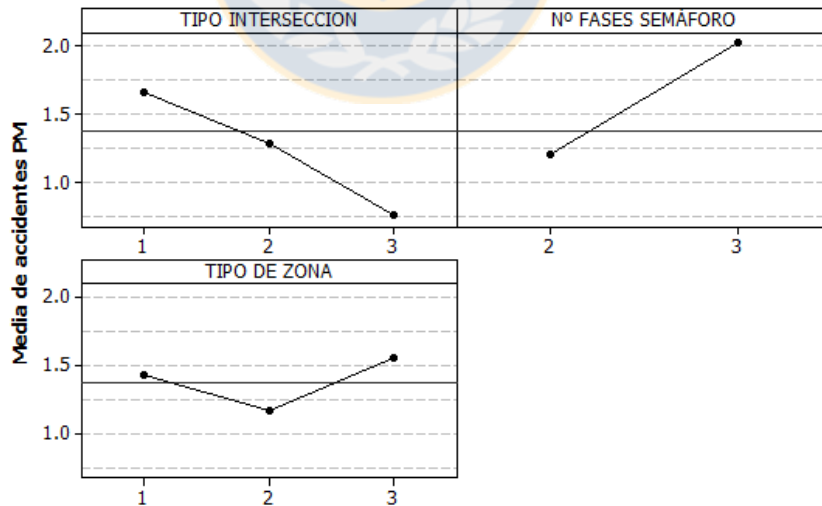


Figura A.5.26 Efectos de otras variables sobre la media de accidentes, base de datos PM

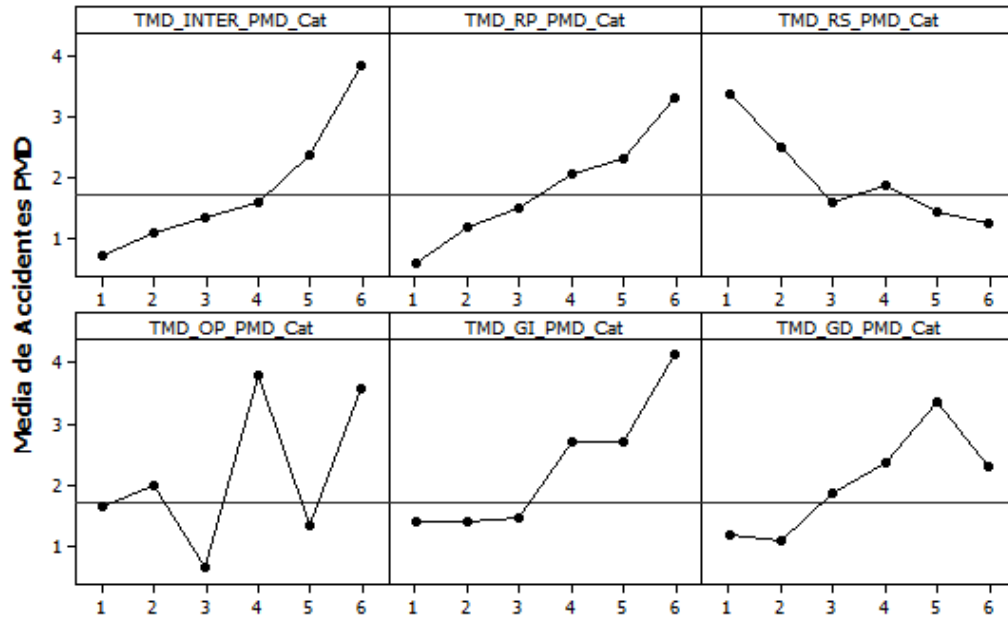


Figura A.5.27 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos PMD

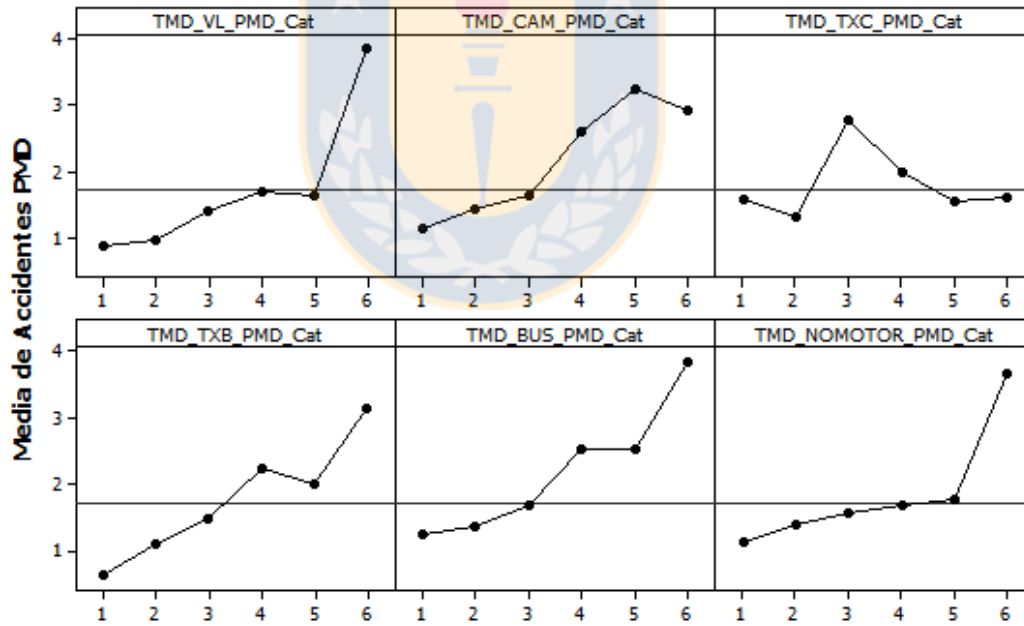


Figura A.5.28 Efectos principales de variables de flujo por tipo de vehículos, base de datos PMD

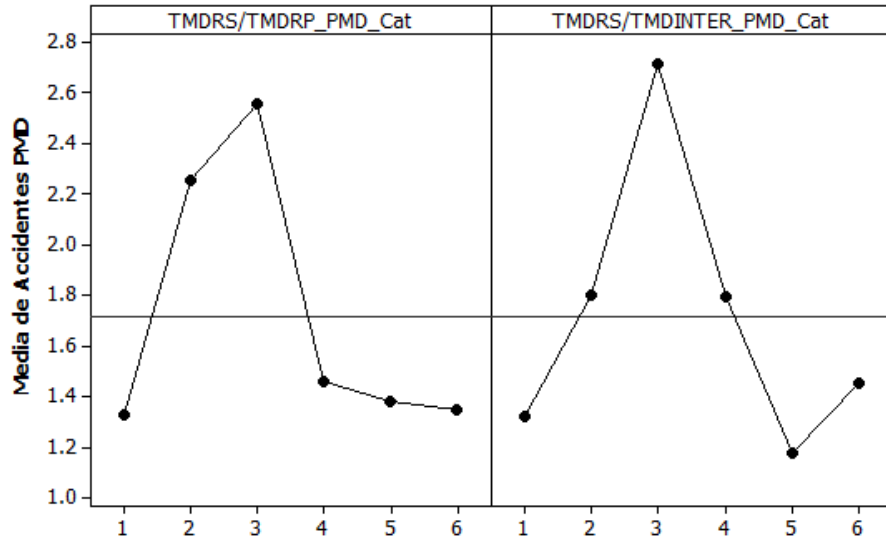


Figura A.5.29 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos PMD

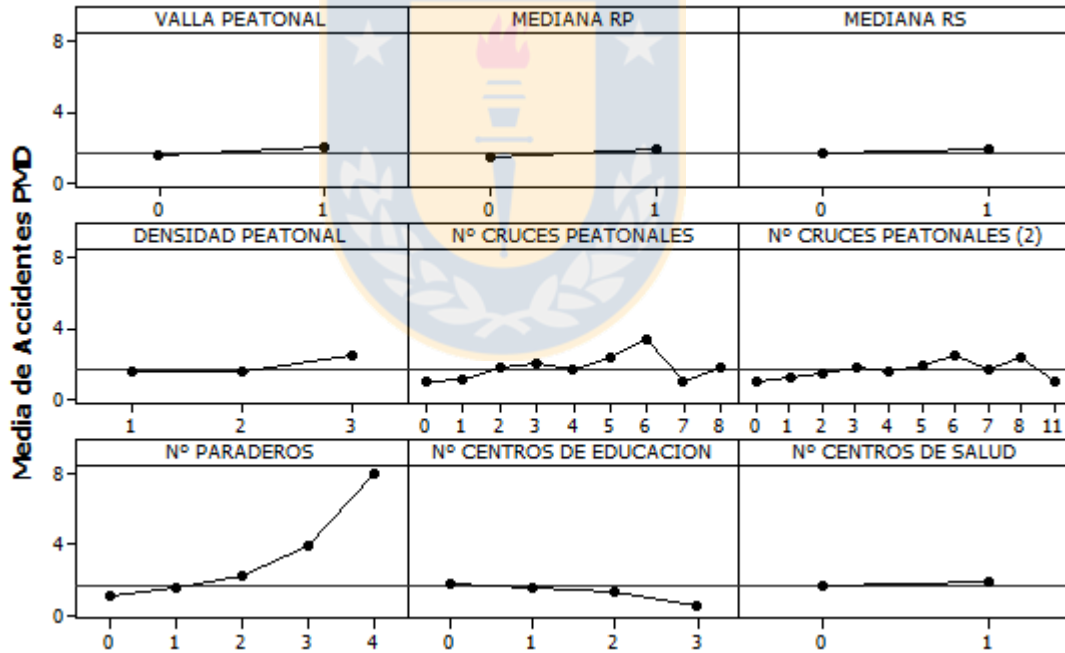


Figura A.5.30 Efectos principales de variables relacionadas al peatón, base de datos PMD

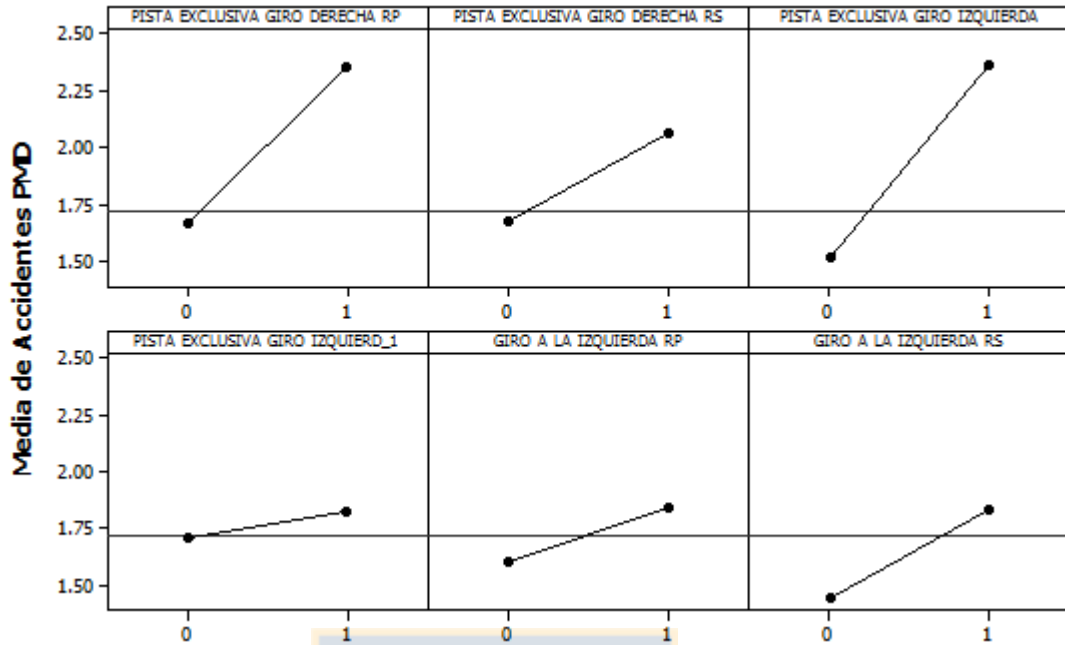


Figura A.5.31 Efectos principales de variables relacionadas a los giros, base de datos PMD

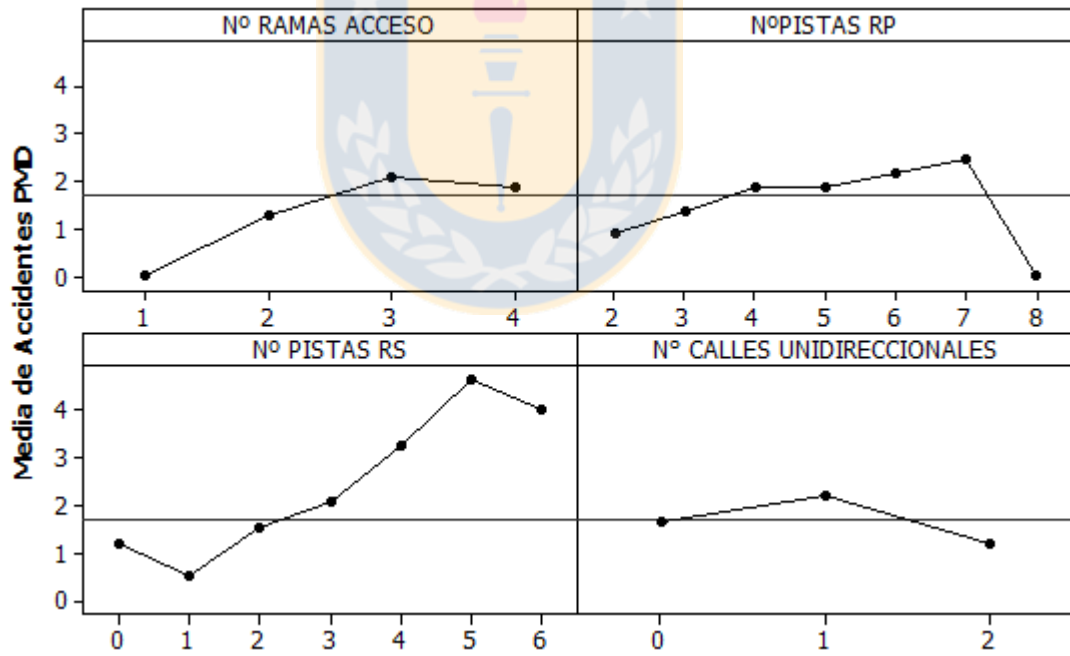


Figura A.5.32 Efectos principales de variables relacionadas al diseño geométrico, base de datos PMD

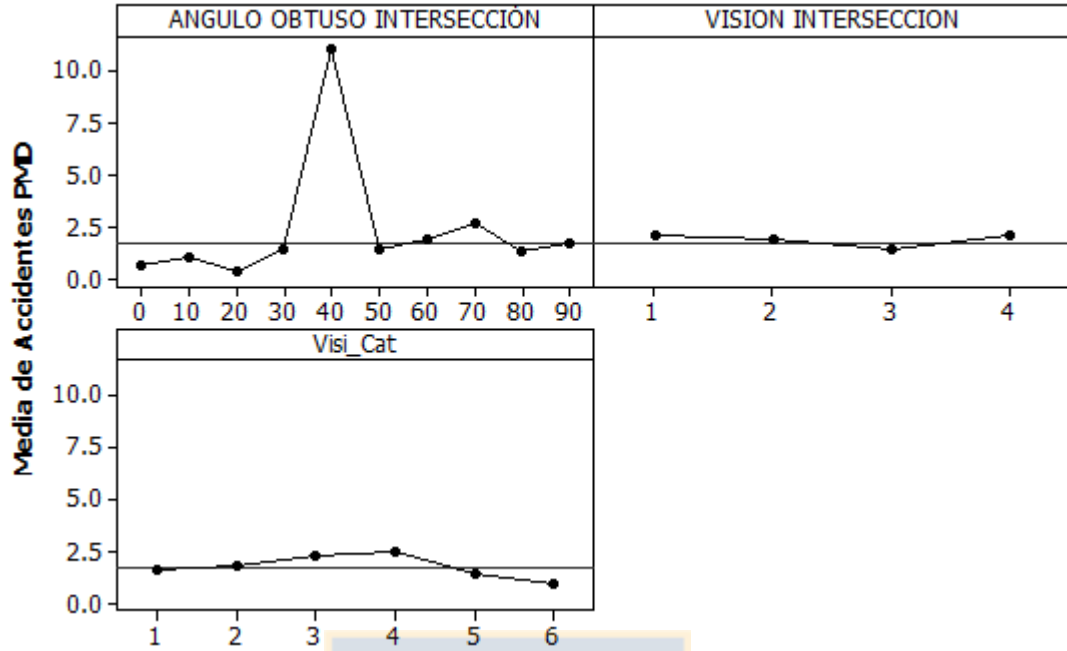


Figura A.5.33 Efectos principales de variables relacionadas a la visibilidad, base de datos PMD

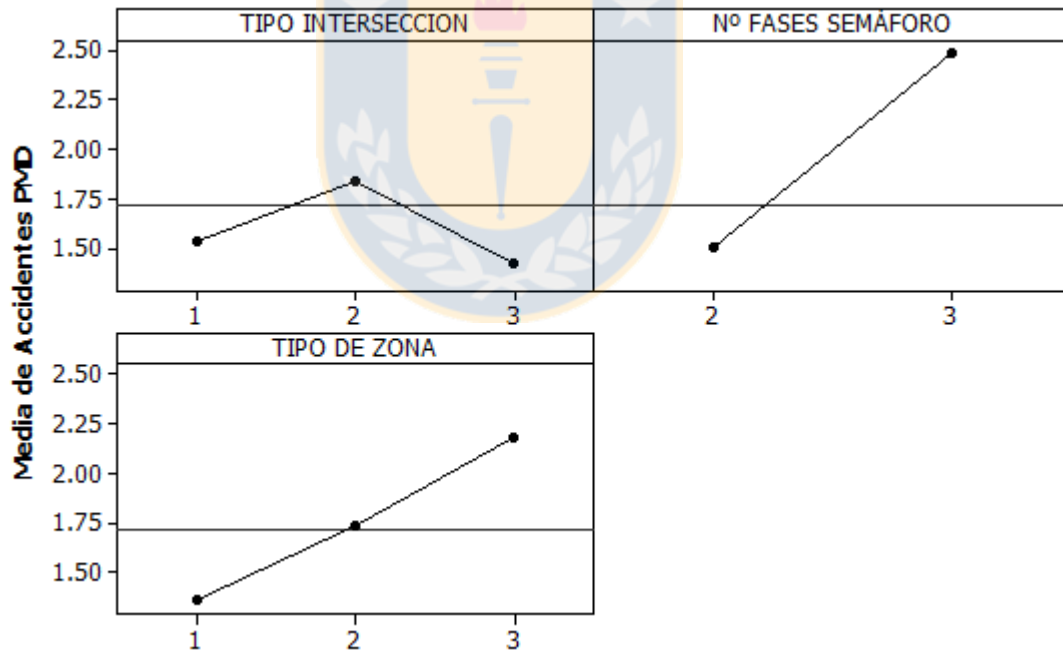


Figura A.5.34 Efectos principales de variables variadas, base de datos PMD

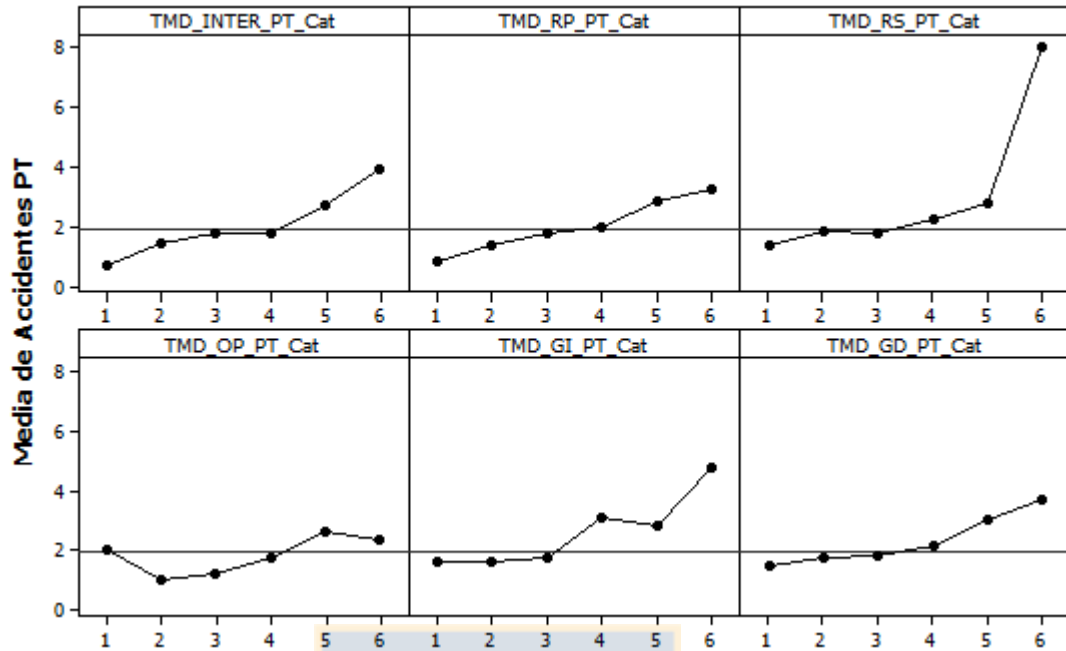


Figura A.5.35 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos PT

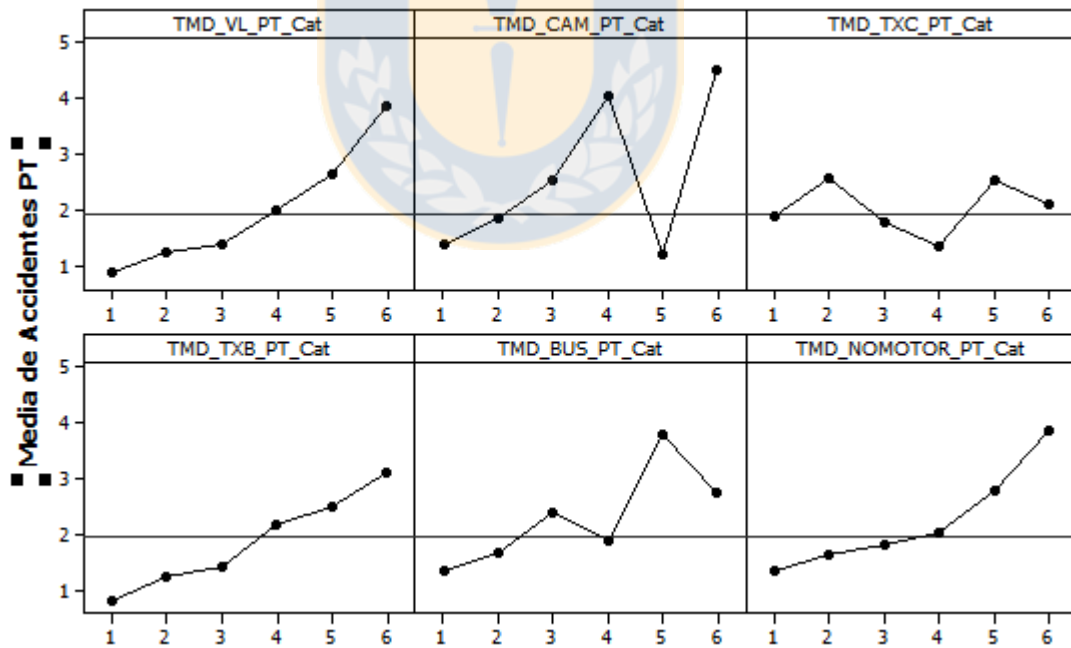


Figura A.5.36 Efectos principales de variables de flujo por tipo de vehículos, base de datos PT

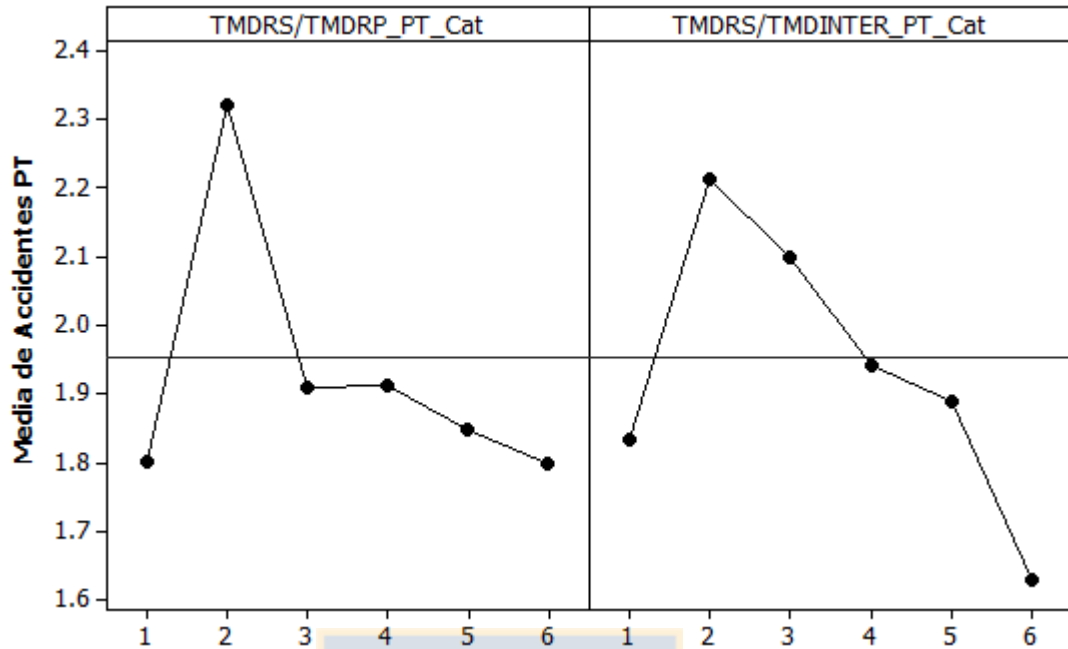


Figura A.5.37 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos PT

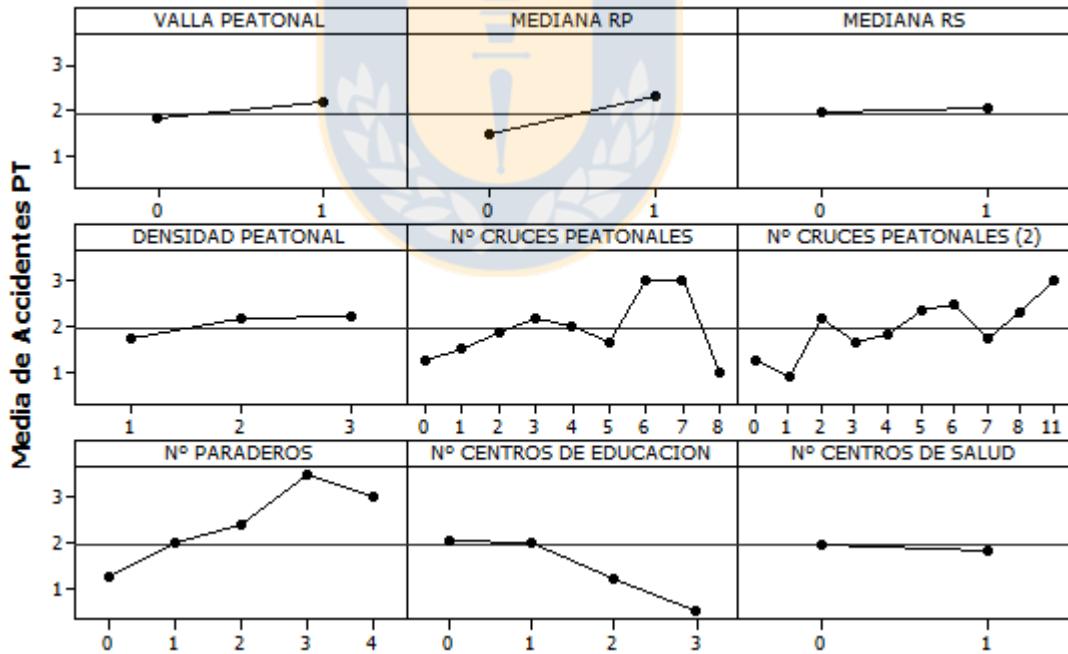


Figura A.5.38 Efectos principales de variables relacionadas al peatón, base de datos PT

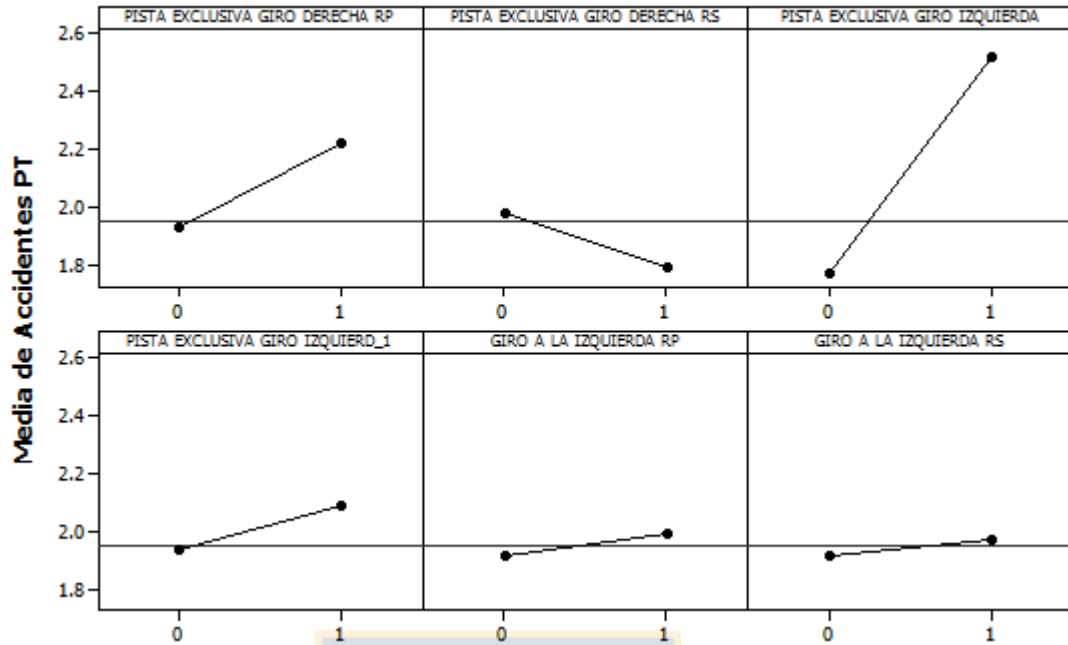


Figura A.5.39 Efectos principales de variables relacionadas a los giros, base de datos PT

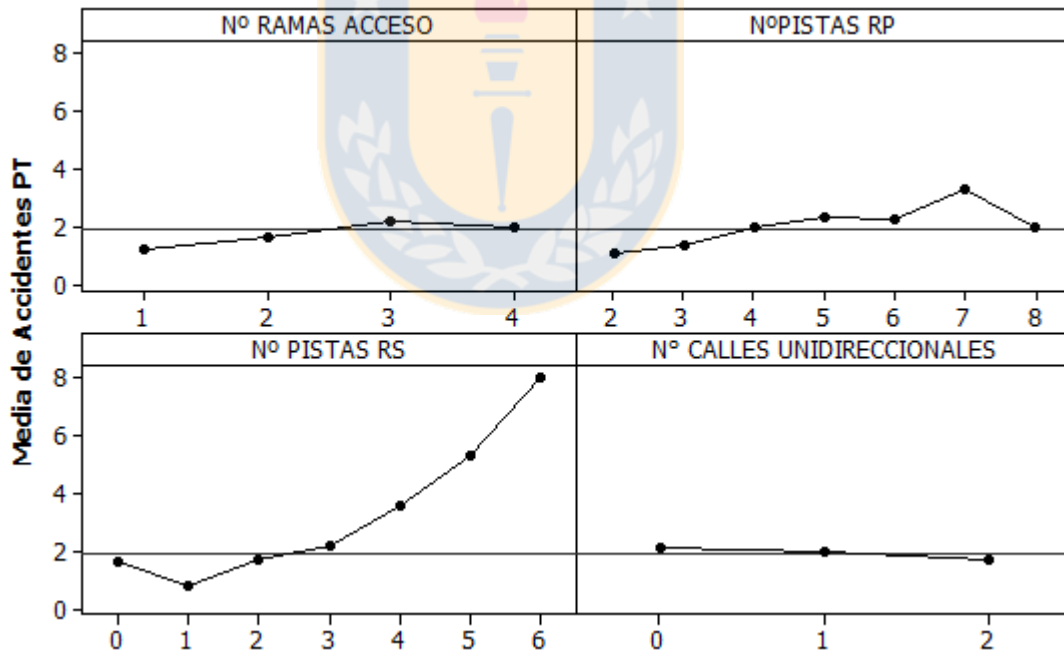


Figura A.5.40 Efectos principales de variables relacionadas al diseño geométrico, base de datos PT

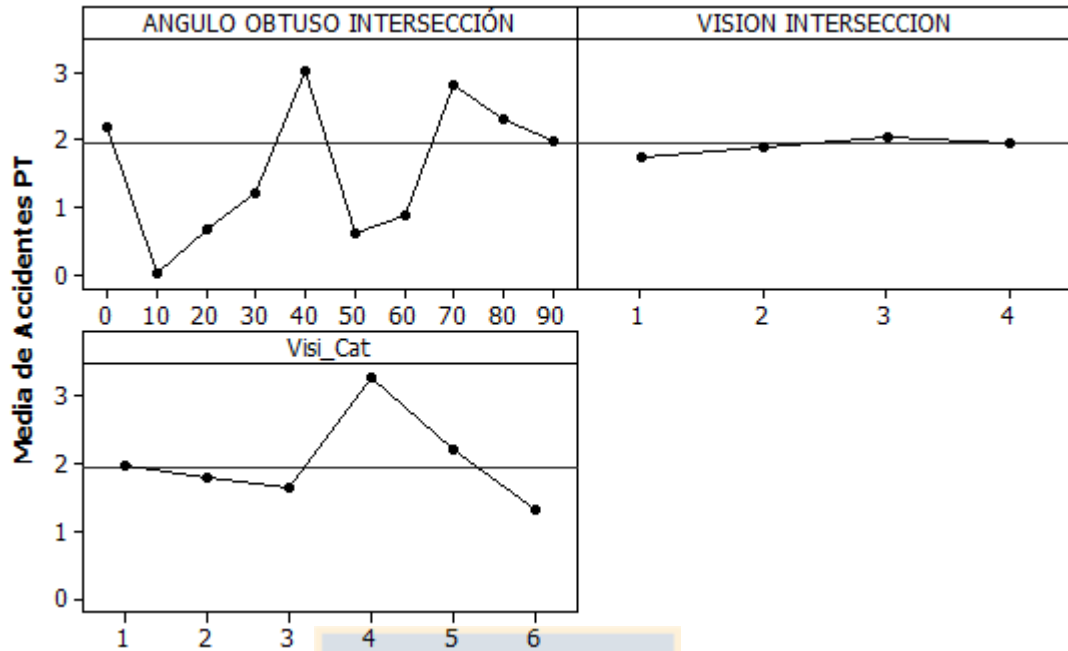


Figura A.5.41 Efectos principales de variables relacionadas a la visibilidad, base de datos PT

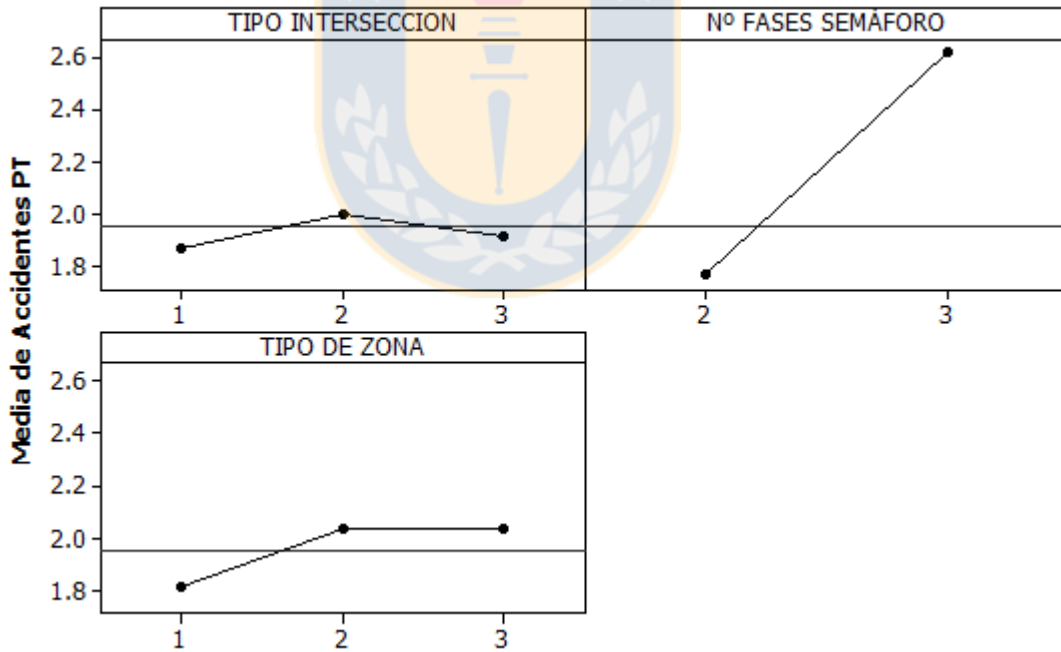


Figura A.5.42 Efectos principales de variables variadas, base de datos PT

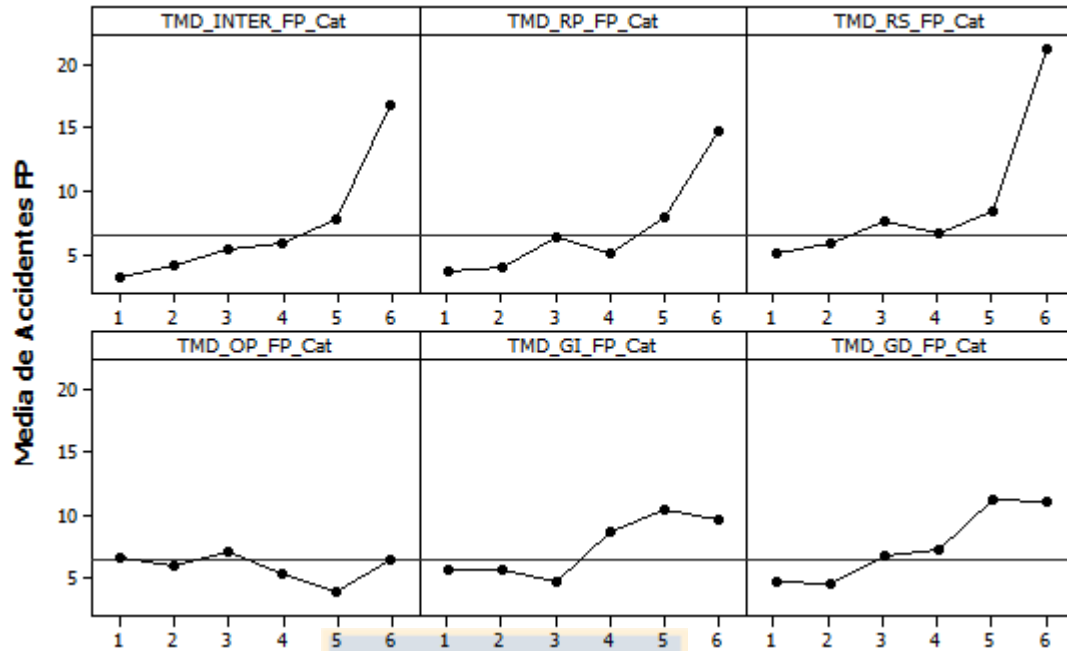


Figura A.5.43 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos FP

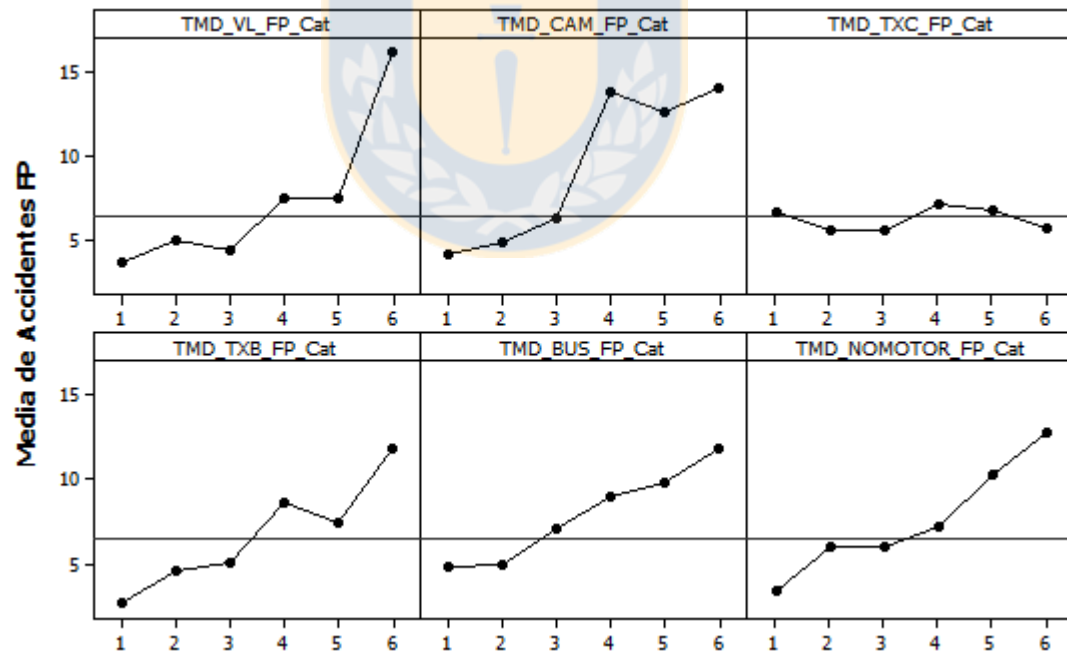


Figura A.5.44 Efectos principales de variables de flujo por tipo de vehículos, base de datos FP

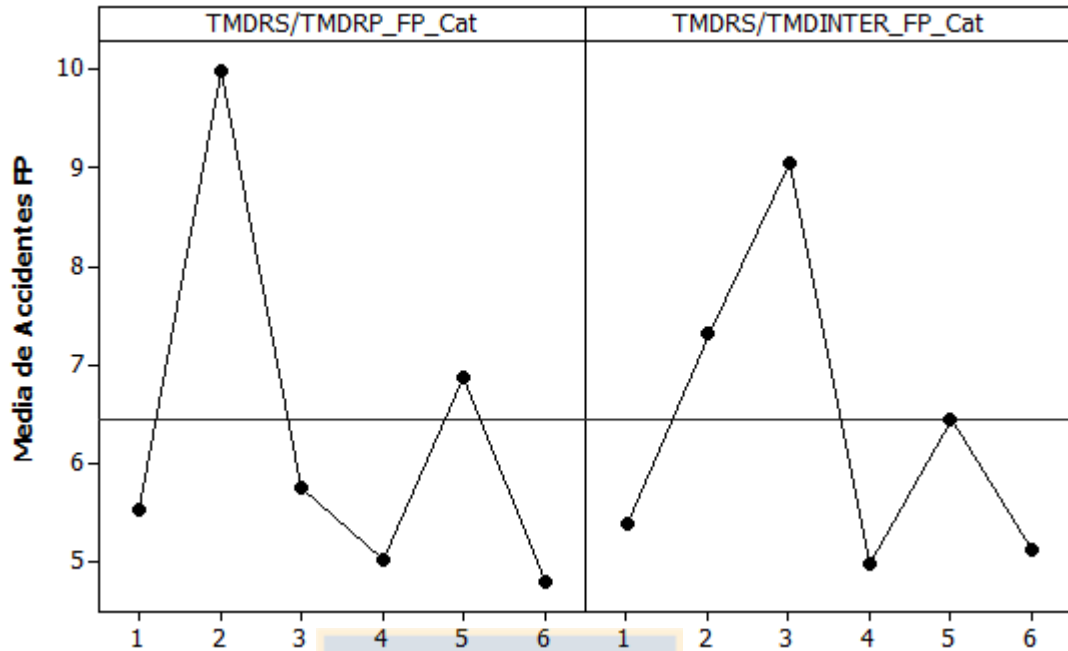


Figura A.5.45 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos FP

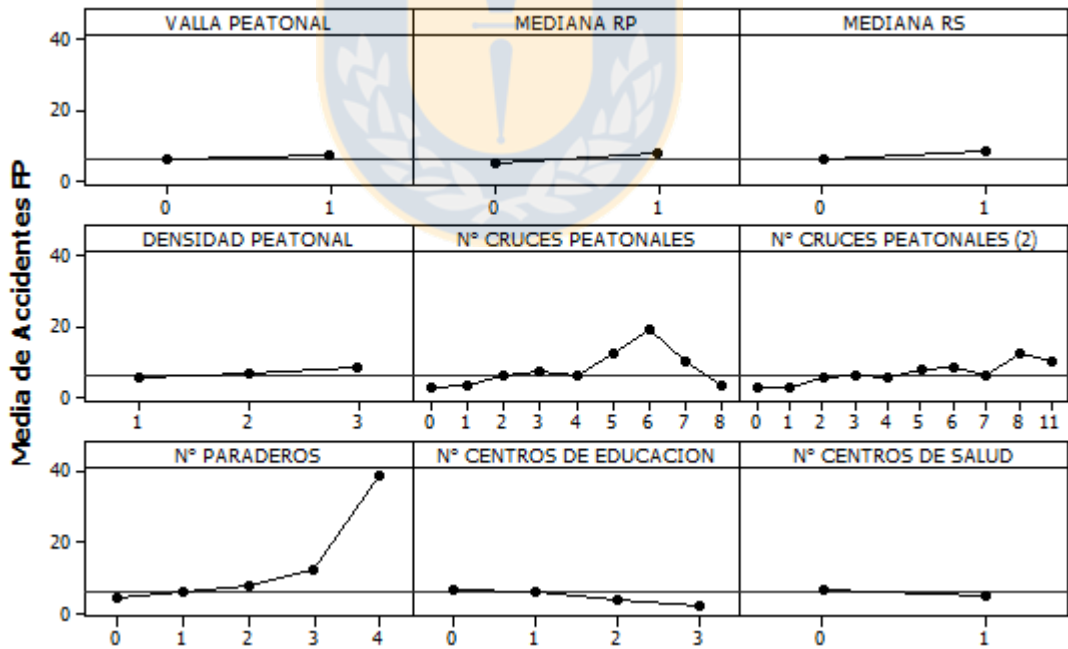


Figura A.5.46 Efectos principales de variables relacionadas al peatón, base de datos FP

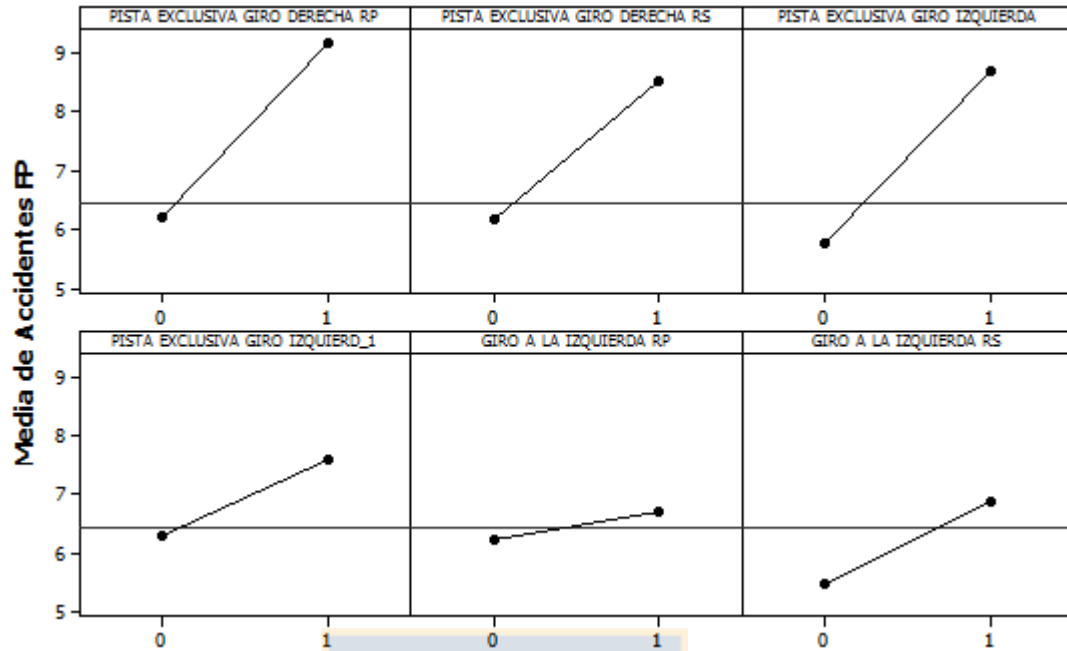


Figura A.5.47 Efectos principales de variables relacionadas a los giros, base de datos FP

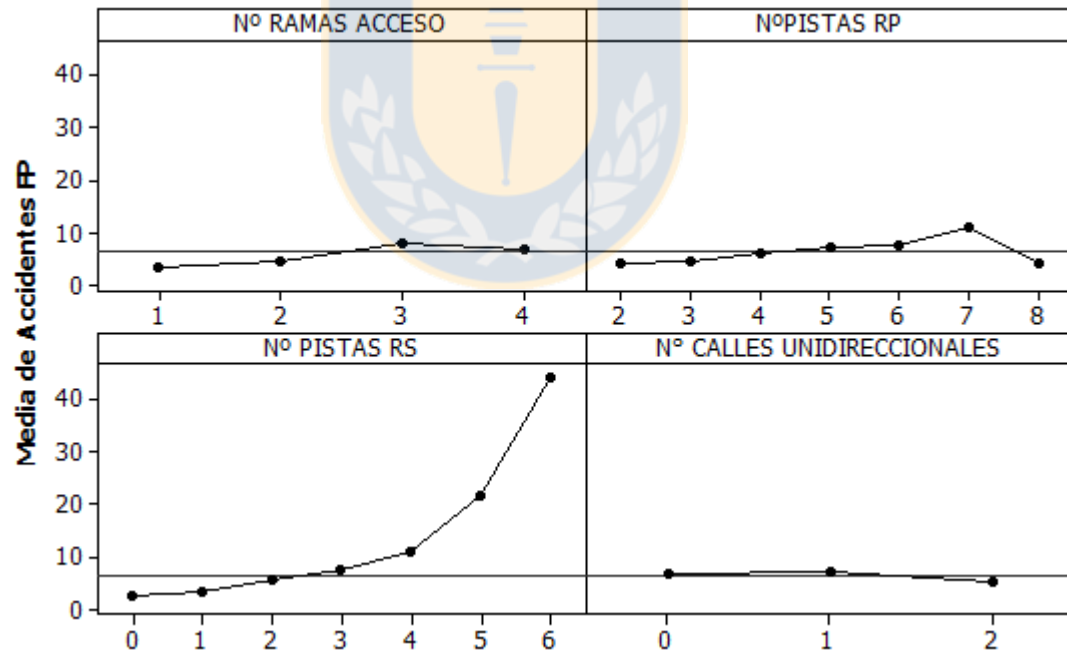


Figura A.5.48 Efectos principales de variables relacionadas al diseño geométrico, base de datos FP

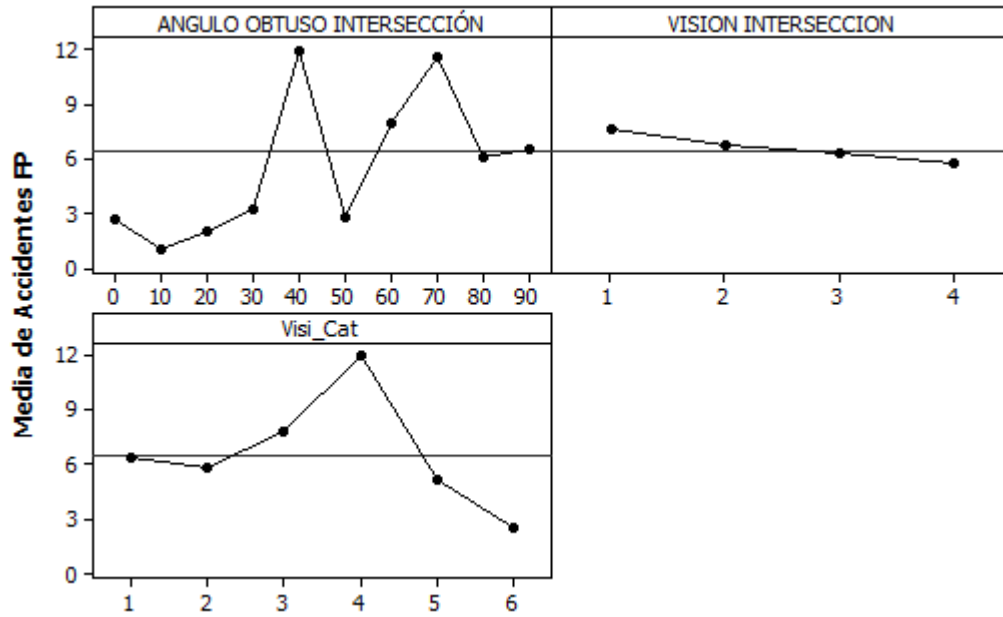


Figura A.5.49 Efectos principales de variables relacionadas a la visibilidad, base de datos FP

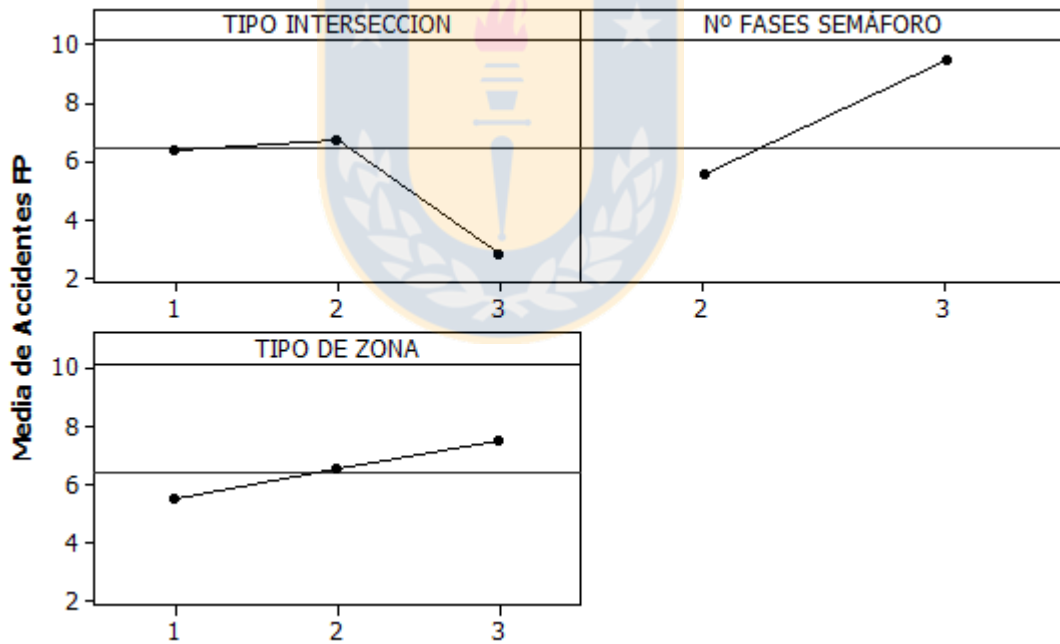


Figura A.5.50 Efectos principales de variables varias, base de datos FP

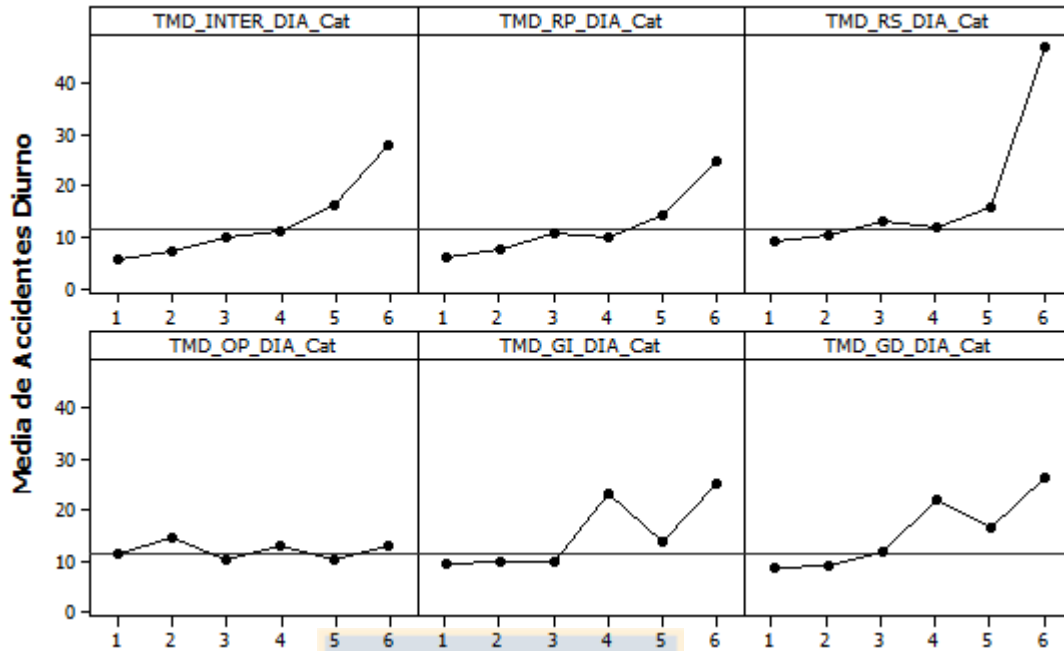


Figura A.5.51 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos diurno

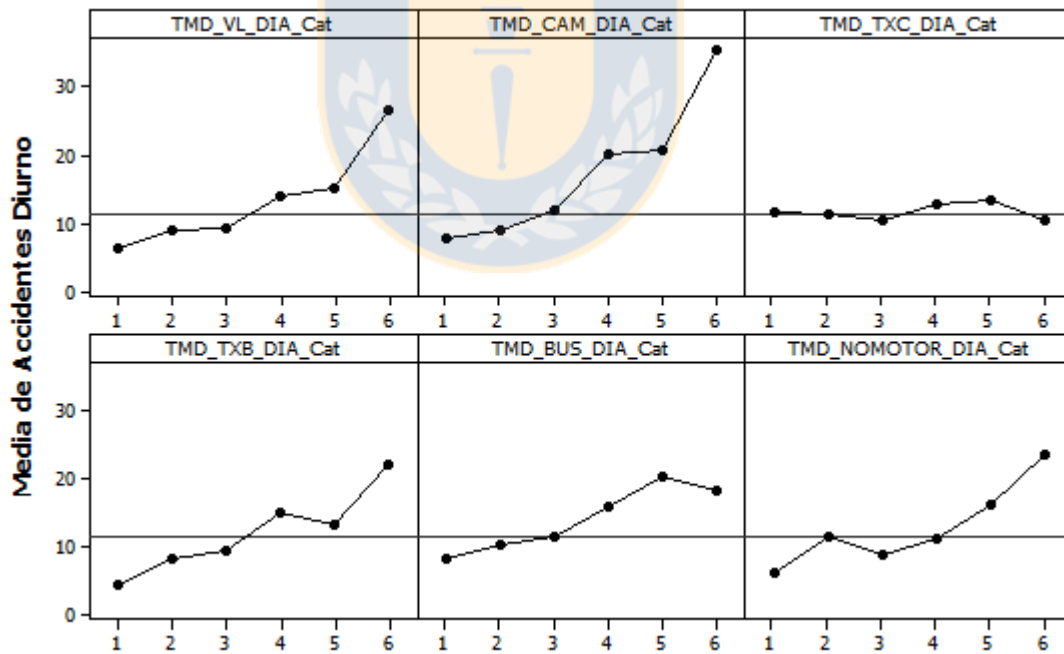


Figura A.5.52 Efectos principales de variables de flujo por tipo de vehículos, base de datos diurno

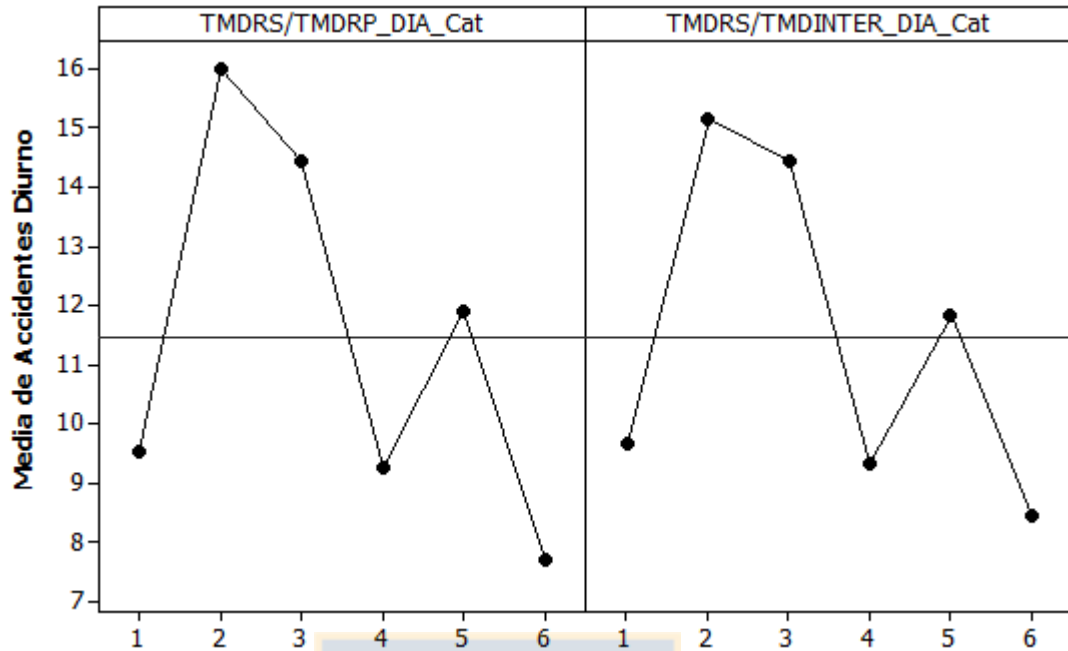


Figura A.5.53 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos diurno

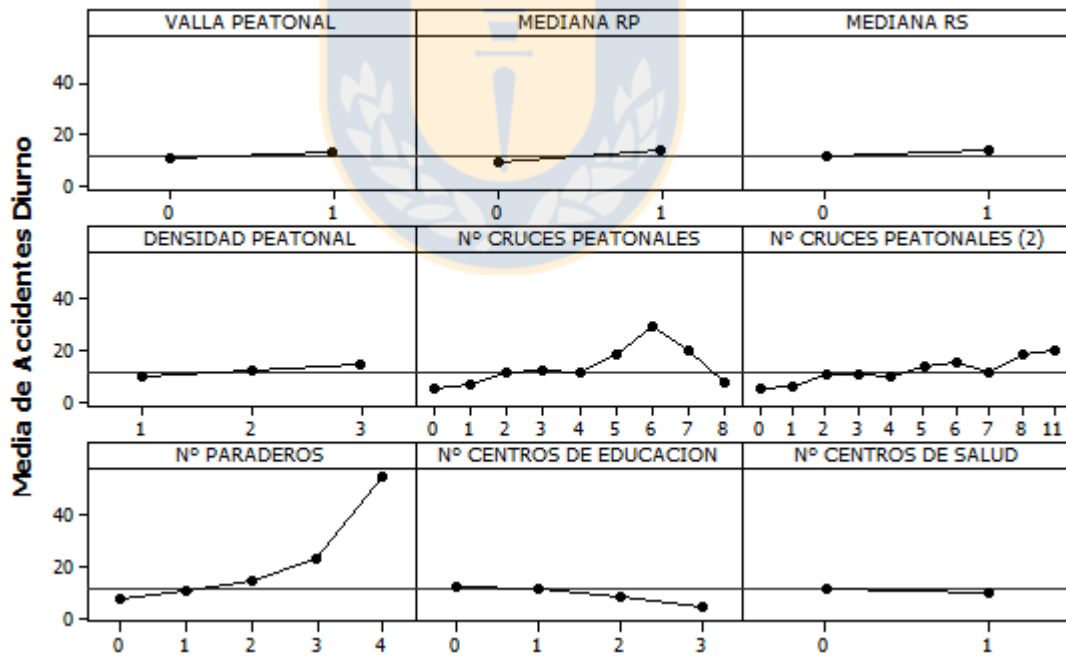


Figura A.5.54 Efectos principales de variables relacionadas al peatón, base de datos diurno

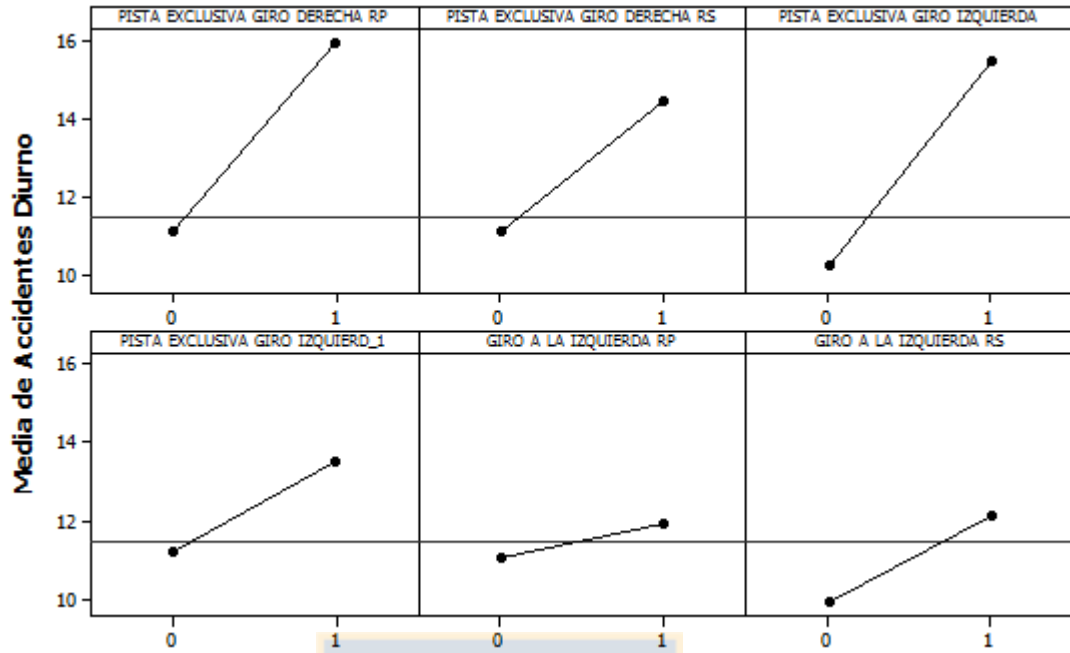


Figura A.5.55 Efectos principales de variables relacionadas a los giros, base de datos diurno

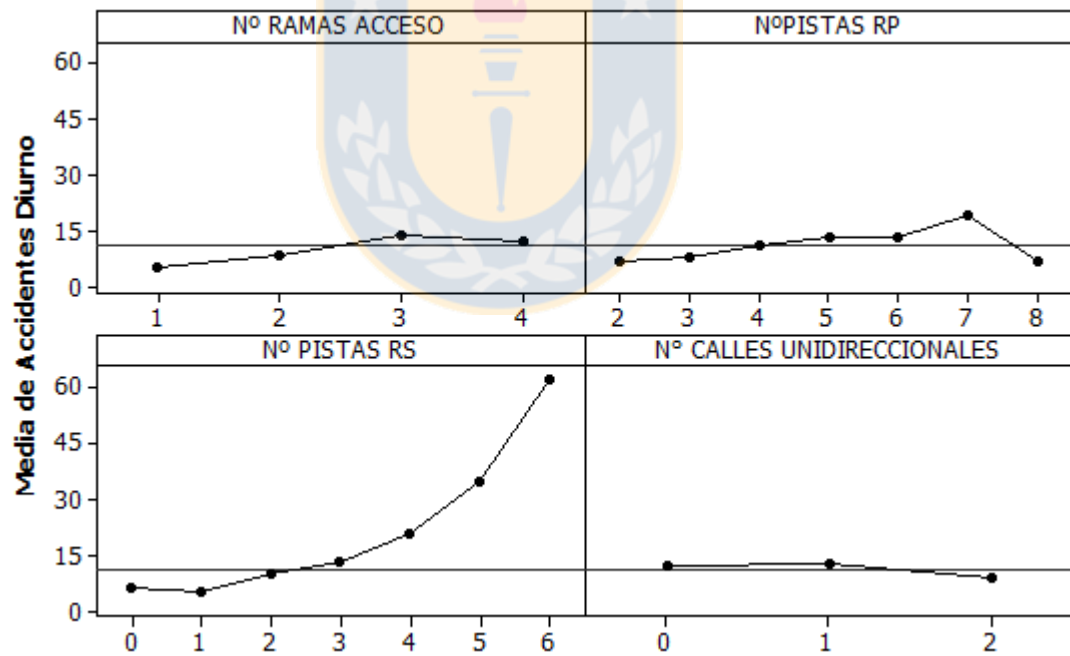


Figura A.5.56 Efectos principales de variables relacionadas al diseño geométrico, base de datos diurno

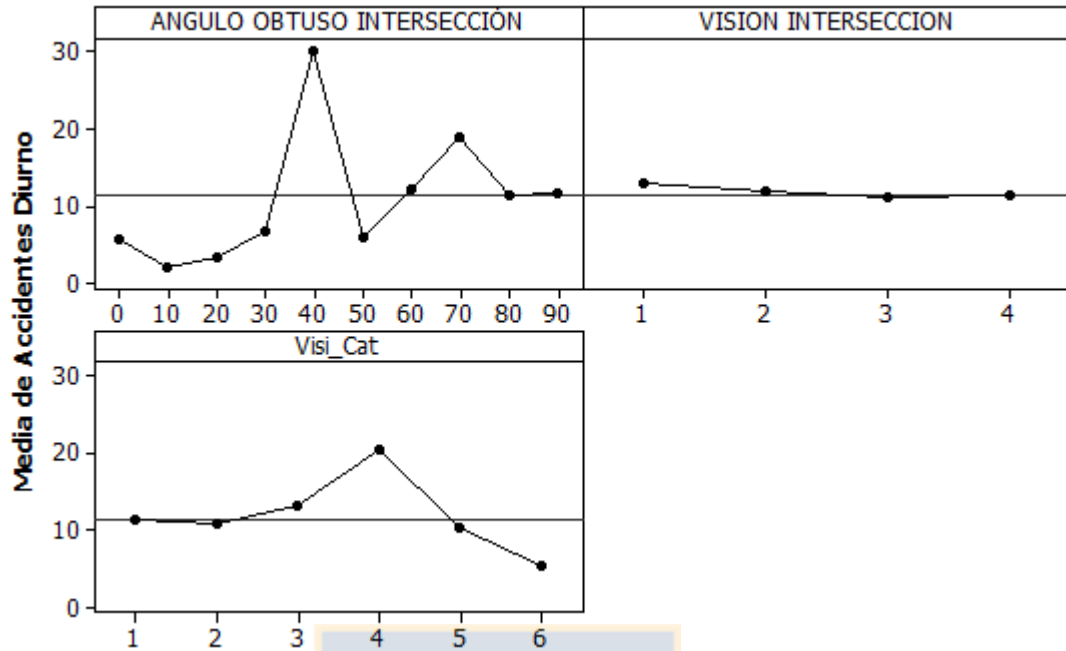


Figura A.5.57 Efectos principales de variables relacionadas a la visibilidad, base de datos diurno

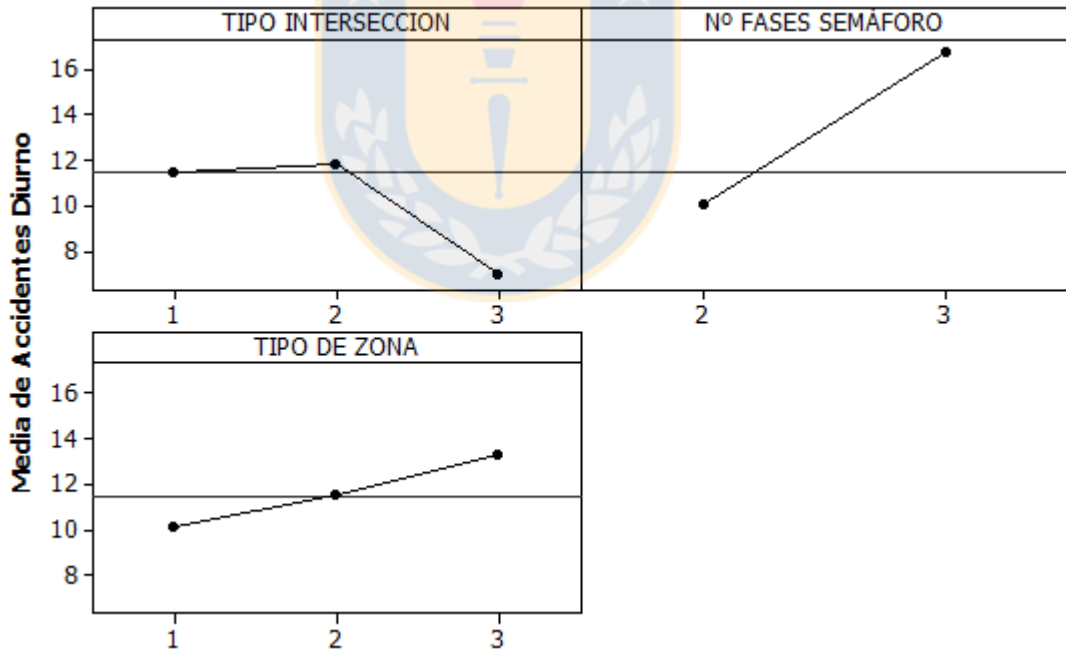


Figura A.5.58 Efectos principales de variables varias, base de datos diurno

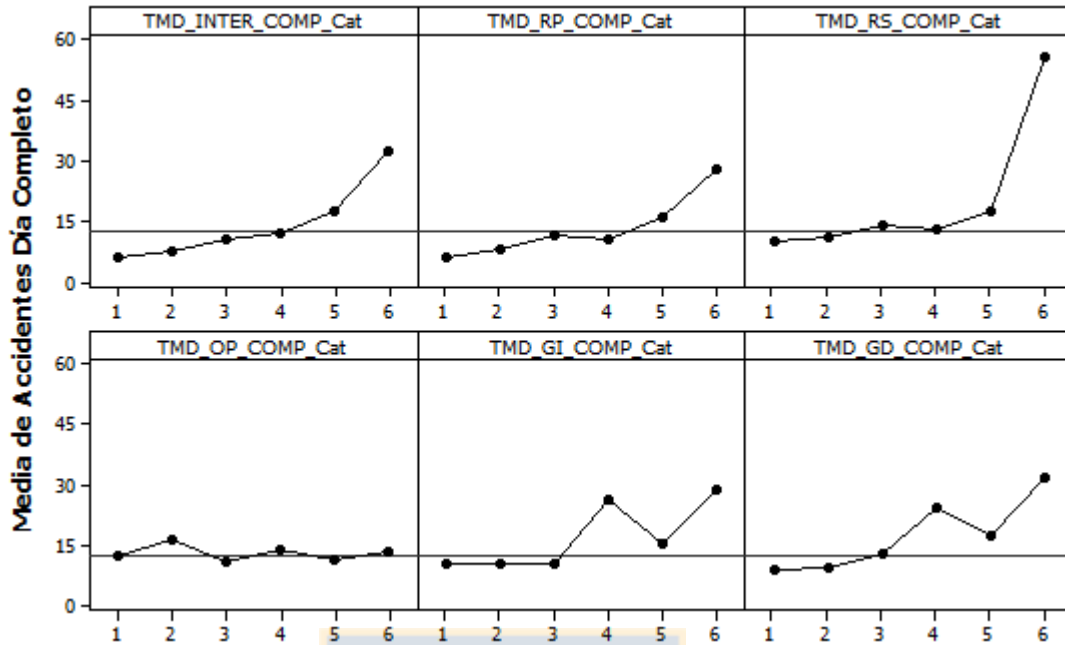


Figura A.5.59 Efectos principales de variables relacionadas al flujo vehicular por movimiento, base de datos completa

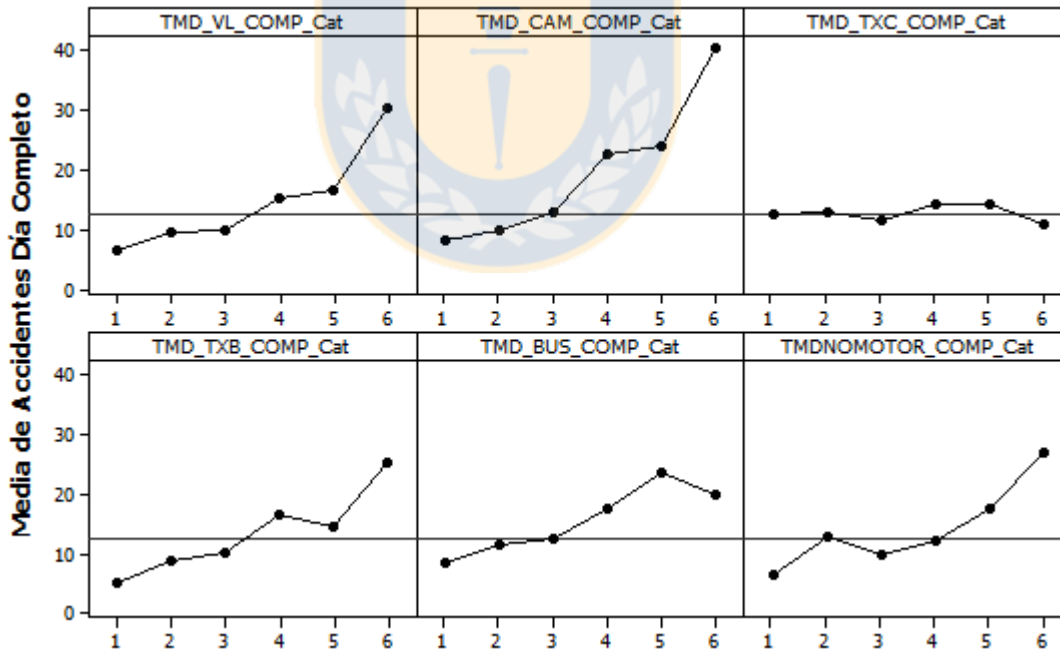


Figura A.5.60 Efectos principales de variables de flujo por tipo de vehículos, base de datos completa

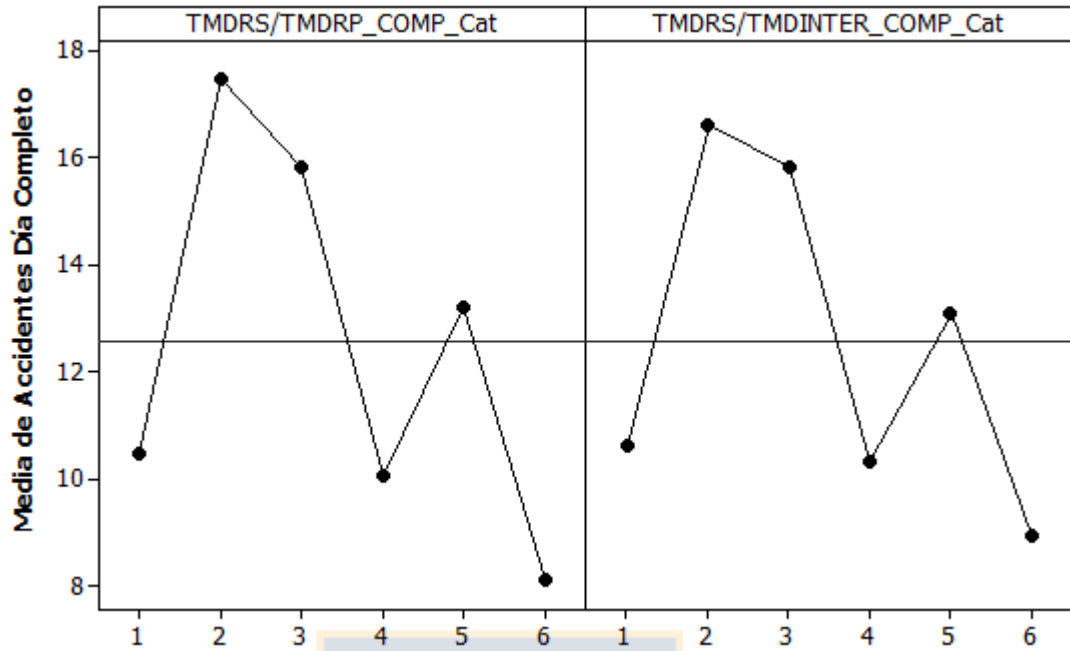


Figura A.5.61 Efectos principales de los cuocientes entre flujo en rama secundaria con principal e intersección, base de datos completa

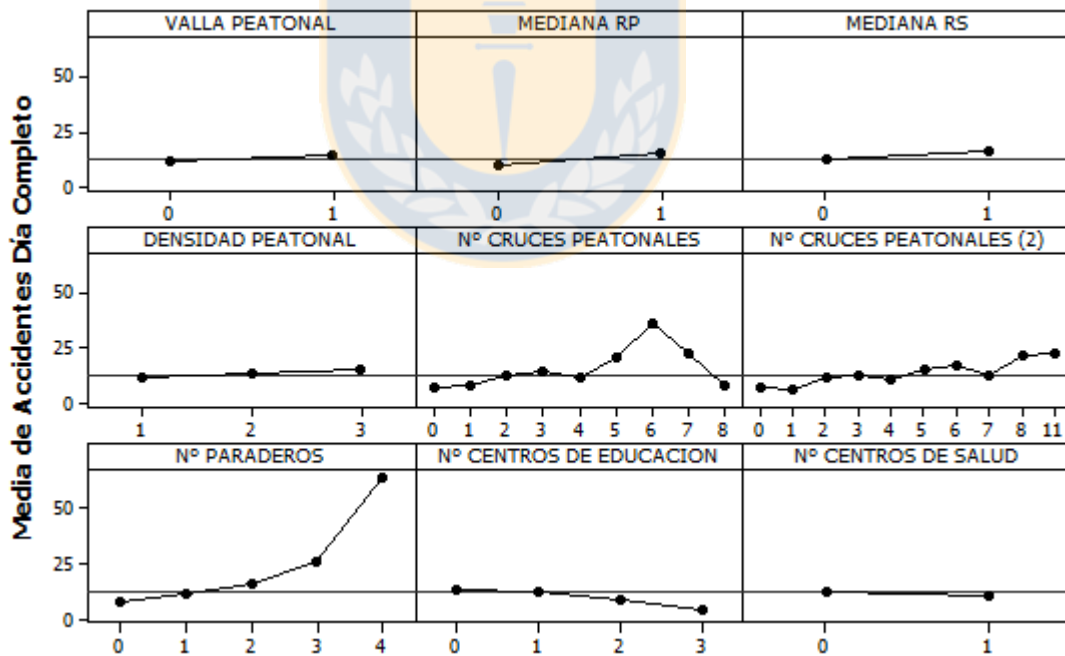


Figura A.5.62 Efectos principales de variables relacionadas al peatón, base de datos completa

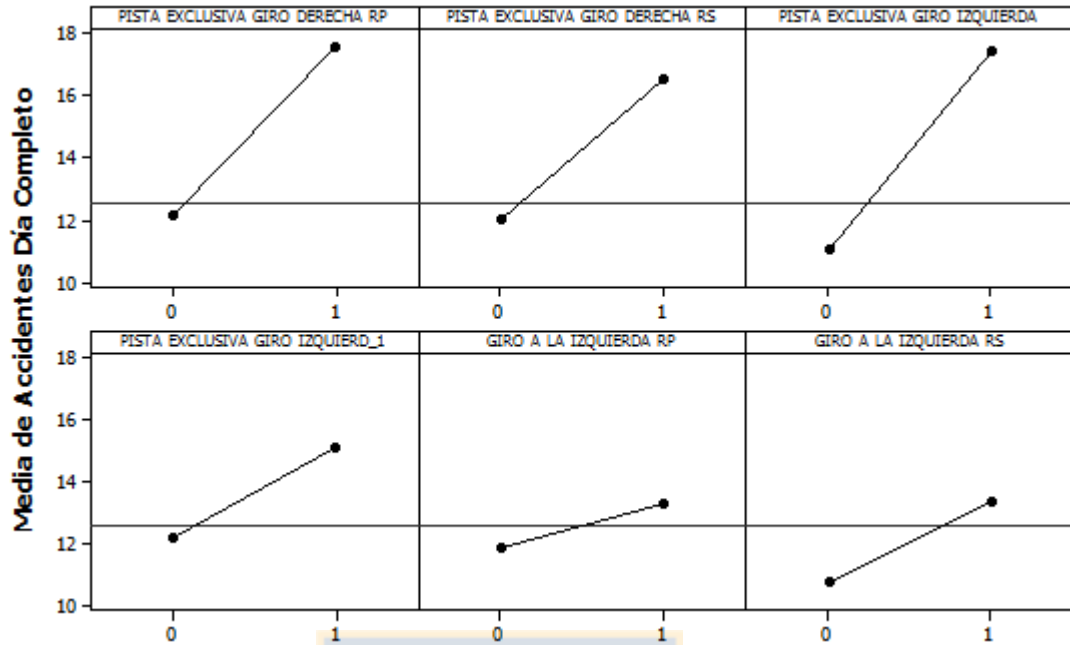


Figura A.5.63 Efectos principales de variables relacionadas a los giros, base de datos completa

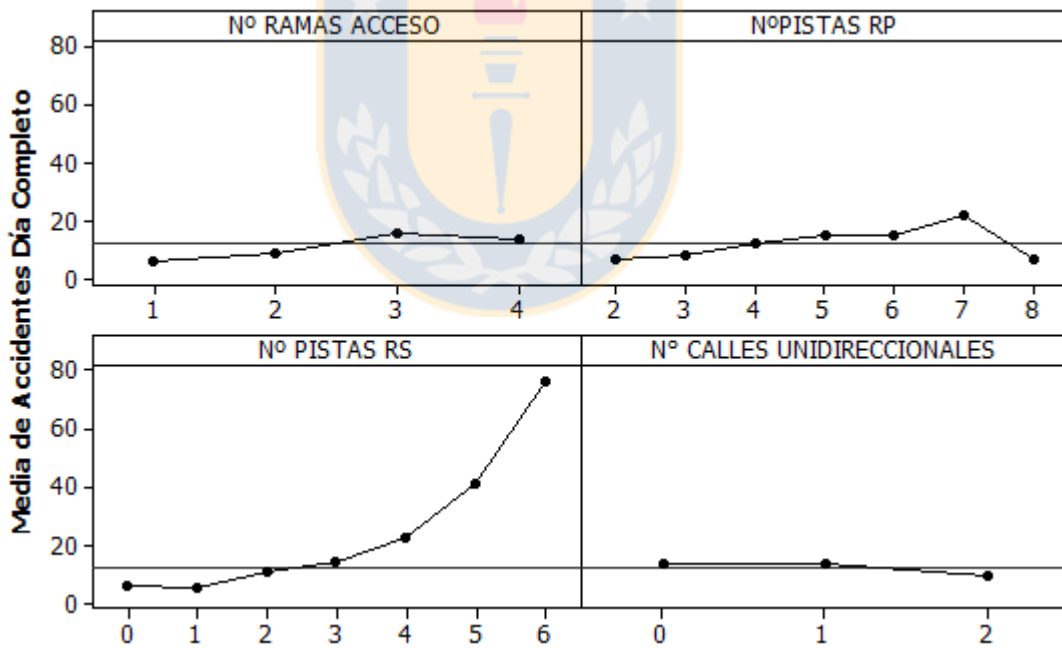


Figura A.5.64 Efectos principales de variables relacionadas al diseño geométrico, base de datos completa

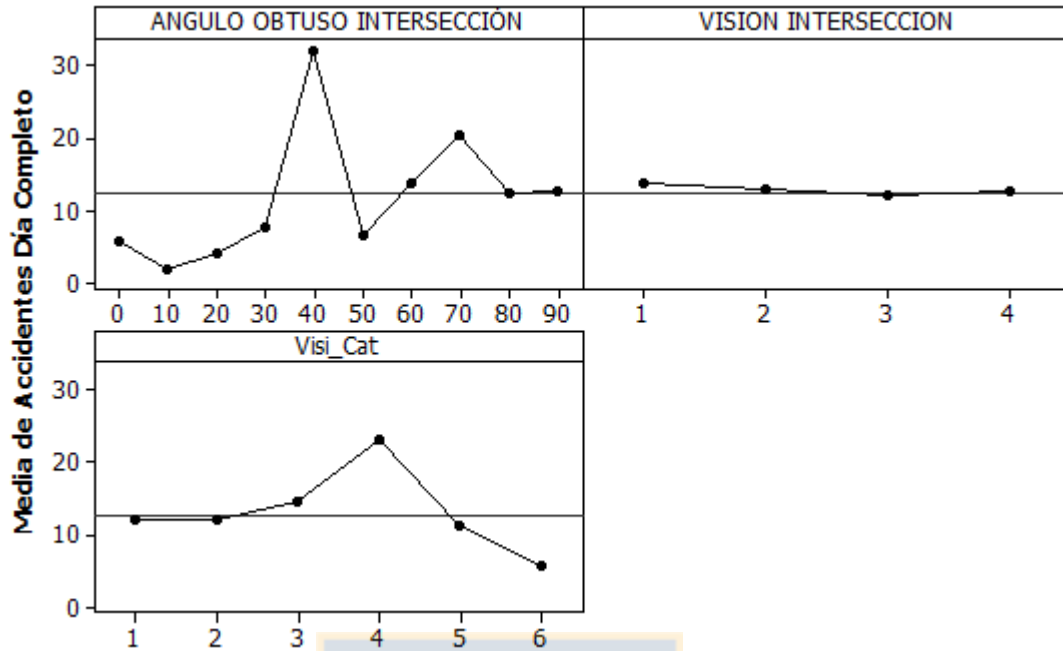


Figura A.5.65 Efectos principales de variables relacionadas a la visibilidad, base de datos completa

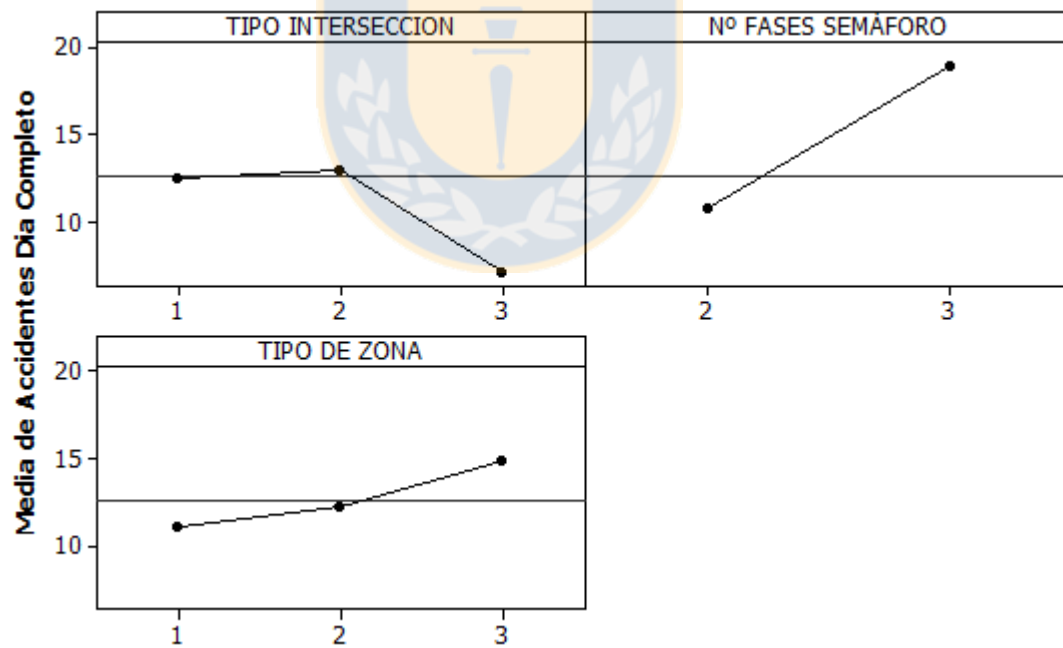


Figura A.5.66 Efectos principales de variables variadas, base de datos completa

ANEXO 5.3 Resultados del análisis de efectos sobre la gran media de accidentes

Tabla A.5.1 Ranking de importancia por grupo de variables según el análisis de efectos de todas las variables por cada grupo sobre la media de accidentes

Variables relacionadas al peaton	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Valla Peatonal	6	8	7	8	8	8
Mediana RP	5	6	4	5	5	5
Mediana RS	4	7	8	7	7	6
Densidad Peatonal	7	4	6	6	6	7
N° Cruces peatonales	2	2	3	3	3	3
N° Cruces peatonales 2	3	5	2	2	2	2
N° Paraderos	1	1	1	1	1	1
N° Educacion	8	3	5	4	4	4
N° Salud	9	9	9	9	9	9
Variables relacionadas a los giros	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Pista Exclusiva GI RP	4	1	1	1	1	1
Pista Exclusiva GI RS	3	4	4	4	4	4
Pista Exclusiva GD RP	1	2	2	2	2	2
Pista Exclusiva GD RS	2	3	3	3	3	3
Giro Izquierda RP	6	6	5	6	6	6
Giro Izquierda RS	5	5	6	5	5	5
Variables relacionadas al diseño geométrico	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
N° de Accesos	2	2	3	3	3	3
N° de Pistas RP	3	3	2	2	2	2
N° de Pistas RS	1	1	1	1	1	1
N° calles unidireccionales	4	4	4	4	4	4
Variables relacionadas a la visibilidad	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Angulo obtuso	2	2	2	3	2	2
Vision Interseccion	3	3	3	2	3	3
Visibilidad	1	1	1	1	1	1
Variables varias	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Tipo interseccion	2	3	3	3	3	3
N° fases semaforo	1	1	1	1	1	1
Tipo de zona	3	2	2	2	2	2

Tabla A.5.2 Observación del comportamiento gráfico de cada variable peatonal en su efecto sobre la media de accidentes por periodo

Variables relacionadas al peatón	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Valla Peatonal	no varia	no varia	aumento	no varia	no varia	no varia
Mediana RP	aumento leve	no varia	aumento	aumento leve	aumento leve	aumento leve
Mediana RS	aumento leve	no varia	no varia	aumento leve	aumento leve	aumento leve
Densidad Peatonal	no varia	aumenta leve	aumenta leve	aumento leve	aumento leve	aumento leve
N° Cruces peatonales	aumento sostenido y disminución	aumento sostenido y disminución	tiende a aumentar (extraño)	aumento sostenido y disminución	aumento sostenido y disminución	aumento sostenido y disminución
N° Cruces peatonales refugio	aumento sostenido	tendencia a media	tiende a aumentar (extraño)	aumento sostenido	aumento sostenido	aumento sostenido
N° Paraderos	aumento sostenido	aumento sostenido	aumento	aumento notorio	aumento notorio	aumento notorio
N° Educación	no varia	disminución	disminución (extraño)	disminución (extraño)	disminución (extraño)	disminución (extraño)
N° Salud	no varia	no varia	no varia	no varia	no varia	no varia

Tabla A.5.3 Observación del comportamiento gráfico de cada variable relacionada a los giros en su efecto sobre la media de accidentes por periodo

Variables relacionadas a los giros	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Pista Exclusiva GI RP	aumento considerable	aumento considerable	aumento considerable	aumento considerable	aumento considerable	aumento considerable
Pista Exclusiva GI RS	aumento considerable	aumento leve	aumenta	aumento considerable	aumento considerable	aumento considerable
Pista Exclusiva GD RP	aumento considerable	aumento considerable	aumenta	aumento considerable	aumento considerable	aumento considerable
Pista Exclusiva GD RS	aumento considerable	aumento	disminuye	aumento considerable	aumento considerable	aumento considerable
Giro Izquierda RP	aumento leve	aumento	aumento leve	aumento leve	aumento leve	aumento
Giro Izquierda RS	aumento	aumento considerable	aumento leve	aumento	aumento	aumento

Tabla A.5.4 Observación del comportamiento gráfico de cada variable de diseño geométrico en su efecto sobre la media de accidentes por periodo

Variabes relacionadas al diseño geométrico	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
N° de Accesos	aumento leve	aumento	aumento leve	aumento leve	aumento leve	aumento leve
N° de Pistas RP	aumento y disminución	aumento y disminución	aumento y disminución	aumento leve y disminución	aumento leve y disminución	aumento leve y disminución
N° de Pistas RS	aumento sostenido	aumento considerable	aumento sostenido	aumento sostenido	aumento sostenido	aumento sostenido
N° calles unidireccionales	disminución leve	disminución leve	no varia	no varia	no varia	disminución leve

Tabla A.5.5 Observación del comportamiento gráfico de cada variable de visibilidad en su efecto sobre la media de accidentes por periodo

Variabes relacionada a la visibilidad	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Angulo obtuso	alto entre 40 y 70°	punto alto en 40	puntos altos en 40 y 70°	alto entre 40 y 70°	puntos altos en 40 y 70°	aumento hasta 40° disminución
Visión Intersección	no varia	no varia	no varia	disminución leve	no varia	no varia
Visibilidad Cat.	aumento y disminución	aumento leve y disminución	aumento y disminución	aumento y disminución	aumento y disminución	aumento y disminución

Tabla A.5.6 Observación del comportamiento gráfico de cada variable en su efecto sobre la media de accidentes por periodo

Variabes varias	Punta mañana	Punta medio día	Punta tarde	Fuera de punta	Diurno	Día completo
Tipo intersección	disminución	varia	variación leve	variación leve	disminución	disminución
N° fases semáforo	aumento	aumento considerable	aumento considerable	aumento	aumento	aumento
Tipo de zona	variación leve	aumento	aumento leve	aumento leve	aumento leve	aumento leve

ANEXO 6.1 Resultados de la modelación exploratoria, formas conceptuales y modelos en cada periodo de análisis

Tabla A.6.1 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 1 en los distintos periodos de análisis.

Modelo 1	Punta mañana		Punta medio día		Punta Tarde		Fuera de Punta		Diurno		Día completo	
	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t
Constante	-1.797	-3.399	-1.039	-1.573	-1.579	-2.074	-4.882	-2.961	-8.448	-3.366	-9.917	-3.591
N°Pistas RS	0.357	2.959	0.312	2.079	0.215	1.269	1.431	3.929	1.314	2.268	1.385	2.172
Angulo Cat	0.011	2.073	-0.001	-0.077	0.007	0.965	0.018	1.081	0.052	2.109	0.065	2.392
Densidad Peatonal	0.129	1.157	0.263	1.814	0.334	2.039	1.074	3.088	1.722	3.280	1.523	2.637
Visibilidad Cat	-0.005	-0.089	-0.027	-0.354	-0.014	-0.161	-0.273	-1.489	-0.363	-1.313	-0.311	-1.021
TMD Intersección	0.000	1.811	0.000	3.245	0.000	3.015	0.000	3.419	0.000	4.241	0.000	4.520
TMD Giro Derecha	0.000	0.715	0.001	1.435	0.000	1.332	0.000	0.367	0.000	1.616	0.000	1.736
TMD Camiones	0.004	4.003	0.001	1.298	0.001	0.701	0.002	3.830	0.001	2.992	0.001	3.103
TMD Taxi Buses	0.001	2.480	0.001	1.012	0.001	1.891	0.000	1.685	0.001	2.508	0.001	2.409

Tabla A.6.2 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 2 en los distintos periodos de análisis.

Modelo 2	Punta mañana		Punta medio día		Punta Tarde		Fuera de Punta		Diurno		Día completo	
	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t
Constante	-1.776	-3.363	-1.017	-1.538	-1.561	-2.051	-5.056	-3.070	-8.501	-3.382	-10.024	-3.628
N°Pistas RS	0.363	3.010	0.314	2.085	0.216	1.274	1.432	3.947	1.352	2.334	1.430	2.245
Angulo Cat	0.011	2.047	0.000	-0.055	0.007	0.973	0.019	1.172	0.053	2.136	0.066	2.428
Densidad Peatonal	0.133	1.187	0.274	1.882	0.345	2.103	1.121	3.229	1.804	3.427	1.622	2.804
Visibilidad Cat	-0.006	-0.094	-0.028	-0.368	-0.017	-0.195	-0.275	-1.505	-0.376	-1.359	-0.326	-1.072
TMD Giro Derecha	0.000	0.723	0.001	1.524	0.000	1.310	0.000	0.353	0.000	1.574	0.000	1.677
TMD Vehiculos Livianos	0.000	1.685	0.000	3.097	0.000	2.951	0.000	3.704	0.000	4.227	0.000	4.575
TMD Camiones	0.004	4.311	0.001	2.145	0.001	1.199	0.002	4.533	0.001	3.835	0.001	3.958
TMD Taxi Buses	0.001	3.244	0.001	1.925	0.001	2.815	0.001	2.787	0.001	3.991	0.001	3.936

Tabla A.6.3 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 3 en los distintos periodos de análisis.

Modelo 3	Punta mañana		Punta medio día		Punta Tarde		Fuera de Punta		Diurno		Día completo	
	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t
Constante	-1.823	-3.460	-1.103	-1.671	-1.639	-2.153	-4.911	-2.987	-8.658	-3.444	-10.166	-3.672
NºPistas RS	0.398	3.747	0.404	2.961	0.308	1.994	1.486	4.490	1.715	3.265	1.859	3.214
Angulo Cat	0.011	2.006	-0.001	-0.222	0.006	0.847	0.017	1.055	0.048	1.955	0.061	2.225
Densidad Peatonal	0.138	1.240	0.296	2.061	0.352	2.158	1.093	3.179	1.821	3.482	1.640	2.848
Visibilidad Cat	-0.008	-0.138	-0.032	-0.415	-0.022	-0.251	-0.279	-1.530	-0.394	-1.424	-0.347	-1.140
TMD Intersección	0.000	1.918	0.000	3.738	0.000	3.701	0.000	3.586	0.000	4.766	0.000	5.079
TMD Camiones	0.004	3.971	0.001	1.104	0.000	0.494	0.002	3.819	0.001	2.817	0.001	2.912
TMD Taxi Buses	0.001	2.534	0.000	0.868	0.001	1.658	0.000	1.664	0.001	2.315	0.001	2.199

Tabla A.6.4 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 4 en los distintos periodos de análisis.

Modelo 4	Punta mañana		Punta medio día		Punta Tarde		Fuera de Punta		Diurno		Día completo	
	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t
Constante	-1.840	-3.929	-1.164	-2.011	-1.627	-2.427	-6.257	-4.368	-10.164	-4.615	-11.368	-4.690
NºPistas RS	0.340	2.796	0.297	1.945	0.224	1.294	1.209	3.274	1.103	1.878	1.223	1.892
Angulo Cat	0.010	1.938	0.000	-0.031	0.008	1.137	0.018	1.120	0.055	2.286	0.069	2.589
Densidad Peatonal	0.134	1.195	0.268	1.839	0.332	2.030	1.131	3.255	1.779	3.384	1.567	2.709
NºCruces peatonales (2)	0.039	0.704	0.020	0.286	-0.024	-0.298	0.325	1.893	0.294	1.127	0.208	0.726
TMD Intersección	0.000	1.703	0.000	3.193	0.000	3.025	0.000	3.120	0.000	4.003	0.000	4.341
TMD Giro Derecha	0.000	0.631	0.001	1.426	0.000	1.352	0.000	0.402	0.000	1.653	0.000	1.772
TMD Camiones	0.004	4.026	0.001	1.296	0.001	0.674	0.002	3.914	0.001	3.050	0.001	3.129
TMD Taxi Buses	0.001	2.512	0.001	1.005	0.001	1.880	0.000	1.768	0.001	2.545	0.001	2.429

Tabla A.6.5 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 5 en los distintos periodos de análisis.

Modelo 5	Punta mañana		Punta medio día		Punta Tarde		Fuera de Punta		Diurno		Día completo	
	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t
Constante	-1.724	-3.706	-1.111	-1.936	-1.623	-2.420	-5.975	-4.142	-9.810	-4.446	-11.075	-4.563
NºPistas RS	0.318	2.660	0.262	1.745	0.201	1.183	1.295	3.508	1.122	1.924	1.216	1.896
Angulo Cat	0.011	2.210	0.001	0.099	0.008	1.103	0.027	1.751	0.064	2.752	0.075	2.945
Densidad Peatonal	0.080	0.706	0.220	1.506	0.323	1.940	1.029	2.904	1.626	3.045	1.437	2.445
NºParaderos	0.236	2.287	0.248	1.901	0.054	0.355	0.309	0.960	0.599	1.230	0.535	0.999
TMD Intersección	0.000	1.487	0.000	2.962	0.000	2.934	0.000	3.305	0.000	4.040	0.000	4.345
TMD Giro Derecha	0.000	0.816	0.001	1.528	0.000	1.362	0.000	0.552	0.000	1.766	0.000	1.856
TMD Camiones	0.003	3.668	0.001	1.194	0.001	0.674	0.001	3.682	0.001	2.886	0.001	3.018
TMD Taxi Buses	0.001	2.120	0.000	0.722	0.001	1.801	0.000	1.446	0.001	2.233	0.001	2.181

Tabla A.6.6 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 6 en los distintos periodos de análisis.

Modelo 6	Punta mañana		Punta medio día		Punta Tarde		Fuera de Punta		Diurno		Día completo	
	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t
Constante	-1.082	-2.403	-1.582	-2.733	-1.690	-2.547	-6.040	-4.100	-7.622	-3.279	-8.490	-3.271
N°Pistas RP	0.184	3.085	0.245	3.203	0.380	4.371	1.235	6.442	1.949	6.570	2.213	6.683
N°Pistas RS	0.513	5.006	0.529	4.040	0.463	3.084	1.866	5.457	2.415	4.422	2.735	4.485
Densidad Peatonal	0.169	1.256	0.180	1.042	0.317	1.645	0.712	1.623	1.371	2.026	1.178	1.560
N°Calles Unidireccionales	-0.112	-0.743	0.213	1.103	0.190	0.874	1.050	2.157	1.419	1.890	1.573	1.875
N°Paraderos	0.384	3.735	0.417	3.154	0.241	1.608	0.946	2.852	1.652	3.211	1.759	3.062
Visibilidad Cat	-0.033	-0.556	0.016	0.212	-0.026	-0.304	-0.302	-1.580	-0.497	-1.683	-0.494	-1.500

Tabla A.6.7 Resultados de coeficientes y test t de cada variable del modelo de forma conceptual 7 en los distintos periodos de análisis.

Modelo 7	Punta mañana		Punta medio día		Punta Tarde		Fuera de Punta		Diurno		Día completo	
	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t	Coef.	Test t
Constante	0.405	1.675	0.328	1.021	0.263	0.751	1.221	1.508	2.326	1.900	2.165	1.620
TMD Rama Secundaria	0.001	3.240	0.001	3.682	0.001	1.924	0.001	2.694	0.001	3.341	0.001	3.430
TMD Giro Derecha	0.000	1.022	0.001	1.326	0.000	1.235	0.000	1.187	0.000	1.768	0.000	1.854
TMD Camiones	0.004	4.744	0.002	3.493	0.001	1.661	0.002	6.532	0.002	5.176	0.002	5.429
TMD Taxi Buses	0.001	2.232	0.001	1.493	0.001	2.694	0.001	2.629	0.001	3.165	0.001	3.087
TMD Buses	-0.003	-0.736	-0.008	-1.206	0.007	1.456	-0.005	-1.237	-0.005	-1.460	-0.004	-1.303
TMD No Motorizado	0.000	0.058	0.007	2.367	0.003	1.191	0.004	1.912	0.006	3.369	0.006	3.585
TMD RS / TMD RP	-1.141	-2.326	-1.835	-2.884	-0.565	-0.830	-2.548	-1.593	-6.692	-2.622	-7.646	-2.744

ANEXO 6.2 Resultados y salidas de modelaciones de Regresión Poisson

A continuación se presenta los resultados entregados por el software estadístico del modelo base (modelo 8) de regresión Poisson con la variable N° Cruces.

Tabla A.6.8 Modelo inicial con variable N° Cruces peatonales (Máxima Verosimilitud)

Parámetro	Estimado	Error	
		Estándar	Razón de Momios
CONSTANTE	0.739751	0.14172	
N° PISTAS RS	0.11429	0.0275044	1.12108
ANGULO_Cat	0.0049452	0.00135808	1.00496
N° CRUCES PEATONALES (2)_Cat	0.0673397	0.0154501	1.06966
TMD_INTER_DIA	0.0000212609	0.00000201162	1.00002
TMD_CAM_DIA	0.0000410353	0.000015491	1.00004
TMD_TXB_DIA	0.000038156	0.0000122105	1.00004
DENSIDAD PEATONAL=1	-0.432258	0.0535969	0.649042
DENSIDAD PEATONAL=2	-0.252203	0.0550262	0.777087

Tabla A.6.9 Análisis de Desviación

Fuente	Desviación	Gl	Valor-P
Modelo	618.176	8	0.0000
Residuo	687.566	249	0.0000
Total (corr.)	1305.74	257	

Porcentaje de desviación explicado por el modelo = 47.3429

Porcentaje ajustado = 45.9644

Tabla A.6.10 Pruebas de Razón de Verosimilitud

Factor	Chi-Cuadrada	Gl	Valor-P
N° PISTAS RS	16.9943	1	0.0000
ANGULO_Cat	14.0793	1	0.0002
N° CRUCES PEATONALES (2)_Cat	18.9445	1	0.0000
TMD_INTER_DIA	113.41	1	0.0000
TMD_CAM_DIA	6.52892	1	0.0106
TMD_TXB_DIA	9.66289	1	0.0019
DENSIDAD PEATONAL	63.768	2	0.0000

Tabla A.6.11 Análisis de Residuos

	Estimación	Validación
n	258	
CME	302.003	
MAE	4.34553	
MAPE		
ME	0.0220919	

La salida muestra los resultados de ajustar un modelo de regresión de Poisson para describir la relación entre ACC_Diurno y 7 variable(s) independiente(s). La ecuación del modelo ajustado es:

$$\text{ACC_Diurno} = \exp(0.739751 + 0.11429 \text{ N}^\circ \text{ PISTAS RS} + 0.0049452 \text{ ANGULO_Cat} + 0.0673397 \text{ N}^\circ \text{ CRUCES PEATONALES (2)_Cat} + 0.0000212609 \text{ TMD_INTER_DIA} + 0.0000410353 \text{ TMD_CAM_DIA} + 0.000038156 \text{ TMD_TXB_DIA} - 0.432258 \text{ DENSIDAD PEATONAL=1} - 0.252203 \text{ DENSIDAD PEATONAL=2})$$

Como el valor-P de la tabla de Análisis de Desviaciones es menor que 0.05, existe una relación estadísticamente significativa entre las variables, con un nivel de confianza del 95.0%. Además, el valor-P para los residuos es menor que 0.05, indicando que el modelo es significativamente peor que el mejor modelo posible para estos datos con un nivel de confianza del 95.0%.

Esta ventana también muestra que el porcentaje de desviación de ACC_Diurno explicado por el modelo es igual a 47.3429%. Este estadístico es similar al estadístico R-Cuadrada habitual. El porcentaje ajustado, que es más apropiado para comparar modelos con diferente número de variables independientes, es 45.9644%.

Para determinar si el modelo puede ser simplificado, note que el valor-P más alto para las pruebas de verosimilitud es 0.0106, que pertenece a TMD_CAM_DIA. Como el valor-P es menor que 0.05, ese término es estadísticamente significativo al nivel de confianza del 95.0%. Consecuentemente, usted probablemente no quisiera eliminar ninguna variable del modelo.

A continuación se presentan los resultados del modelo analizado eliminando la variable del flujo de camiones (TMD_CAM) ya que resultó como la menos significativa en los resultados anteriores.

Tabla A.6.12 Modelo inicial con variable N° Cruces peatonales sin Flujo Camiones (Máxima Verosimilitud)

Parámetro	Estimado	Error	
		Estandar	Razón de Momios
CONSTANTE	0.719491	0.140994	
N° PISTAS RS	0.115155	0.0273793	1.12205
ANGULO_Cat	0.00485554	0.00135358	1.00487
N° CRUCES PEATONALES (2)_Cat	0.0667541	0.0154665	1.06903
TMD_INTER_DIA	0.0000245535	0.0000015676	1.00002
TMD_TXB_DIA	0.0000299533	0.0000117633	1.00003
DENSIDAD PEATONAL=1	-0.425314	0.0535134	0.653564
DENSIDAD PEATONAL=2	-0.25122	0.0550105	0.777851

Tabla A.6.13 Análisis de Desviación

Fuente	Desviación	Gl	Valor-P
Modelo	611.647	7	0.0000
Residuo	694.094	250	0.0000
Total (corr.)	1305.74	257	

Porcentaje de desviación explicado por el modelo = 46.8429

Porcentaje ajustado = 45.6175

Tabla A.6.14 Pruebas de Razón de Verosimilitud

Factor	Chi-Cuadrada	Gl	Valor-P
N° PISTAS RS	17.4071	1	0.0000
ANGULO_Cat	13.6438	1	0.0002
N° CRUCES PEATONALES (2)_Cat	18.5854	1	0.0000
TMD_INTER_DIA	231.401	1	0.0000
TMD_TXB_DIA	6.4174	1	0.0113
DENSIDAD PEATONAL	61.798	2	0.0000

Tabla A.6.15 Análisis de Residuos

	Estimación	Validación
n	258	
CME	306.472	
MAE	4.44565	
MAPE		
ME	0.0290659	

La salida muestra los resultados de ajustar un modelo de regresión de Poisson para describir la relación entre ACC_Diurno y 6 variable(s) independiente(s). La ecuación del modelo ajustado es:

$$\text{ACC_Diurno} = \exp(0.719491 + 0.115155 \text{ N° PISTAS RS} + 0.00485554 \text{ ANGULO_Cat} + 0.0667541 \text{ N° CRUCES PEATONALES (2)_Cat} + 0.0000245535 \text{ TMD_INTER_DIA} +$$

0.0000299533 TMD_TXB_DIA - 0.425314 DENSIDAD PEATONAL=1 - 0.25122 DENSIDAD PEATONAL=2)

Como el valor-P de la tabla de Análisis de Desviaciones es menor que 0.05, existe una relación estadísticamente significativa entre las variables, con un nivel de confianza del 95.0%. Además, el valor-P para los residuos es menor que 0.05, indicando que el modelo es significativamente peor que el mejor modelo posible para estos datos con un nivel de confianza del 95.0%.

Esta ventana también muestra que el porcentaje de desviación de ACC_Diurno explicado por el modelo es igual a 46.8429%. Este estadístico es similar al estadístico R-Cuadrada habitual. El porcentaje ajustado, que es más apropiado para comparar modelos con diferente número de variables independientes, es 45.6175%.

Para determinar si el modelo puede ser simplificado, note que el valor-P más alto para las pruebas de verosimilitud es 0.0113, que pertenece a TMD_TXB_DIA. Como el valor-P es menor que 0.05, ese término es estadísticamente significativo al nivel de confianza del 95.0%. Consecuentemente, usted probablemente no quisiera eliminar ninguna variable del modelo.

A continuación se presentan los resultados del modelo analizado anteriormente eliminando además del flujo de camiones, el flujo de Taxibuses (TMD_TXB). Esto pues dicha variable en el intento anterior, resultó como aquella de menor significancia estadística.

Tabla A.6.16 Modelo inicial con variable N° Cruces peatonales sin Flujo Camiones y Taxibuses (Máxima Verosimilitud)

Parámetro	Estimado	Error	Razón de Momios
		Estandar	Estimada
CONSTANTE	0.817965	0.13508	
N° PISTAS RS	0.114714	0.0273689	1.12155
ANGULO_Cat	0.00438189	0.00134301	1.00439
N° CRUCES PEATONALES (2)_Cat	0.066943	0.0154359	1.06923
TMD_INTER_DIA	0.0000264269	0.00000137202	1.00003
DENSIDAD PEATONAL=1	-0.44254	0.053172	0.642403
DENSIDAD PEATONAL=2	-0.268305	0.0546111	0.764675

Tabla A.6.17 Análisis de Desviación

Fuente	Desviación	Gl	Valor-P
Modelo	605.23	6	0.0000
Residuo	700.512	251	0.0000
Total (corr.)	1305.74	257	

Porcentaje de desviación explicado por el modelo = 46.3514

Porcentaje ajustado = 45.2792

Tabla A.6.18 Pruebas de Razón de Verosimilitud

Factor	Chi-Cuadrada	Gl	Valor-P
Nº PISTAS RS	17.2855	1	0.0000
ANGULO_Cat	11.2517	1	0.0008
Nº CRUCES PEATONALES (2)_Cat	18.7627	1	0.0000
TMD_INTER_DIA	355.721	1	0.0000
DENSIDAD PEATONAL	67.0907	2	0.0000

Tabla A.6.19 Análisis de Residuos

	Estimación	Validación
n	258	
CME	301.994	
MAE	4.41678	
MAPE		
ME	0.0376017	

La salida muestra los resultados de ajustar un modelo de regresión de Poisson para describir la relación entre ACC_Diurno y 5 variable(s) independiente(s). La ecuación del modelo ajustado es :

$$\text{ACC_Diurno} = \exp(0.817965 + 0.114714 \text{ N}^\circ \text{ PISTAS RS} + 0.00438189 \text{ ANGULO_Cat} + 0.066943 \text{ N}^\circ \text{ CRUCES PEATONALES (2)_Cat} + 0.0000264269 \text{ TMD_INTER_DIA} - 0.44254 \text{ DENSIDAD PEATONAL=1} - 0.268305 \text{ DENSIDAD PEATONAL=2})$$

Como el valor-P de la tabla de Análisis de Desviaciones es menor que 0.05, existe una relación estadísticamente significativa entre las variables, con un nivel de confianza del 95.0%. Además, el valor-P para los residuos es menor que 0.05, indicando que el modelo es significativamente peor que el mejor modelo posible para estos datos con un nivel de confianza del 95.0%.

Esta ventana también muestra que el porcentaje de desviación de ACC_Diurno explicado por el modelo es igual a 46.3514%. Este estadístico es similar al estadístico R-Cuadrada habitual. El

porcentaje ajustado, que es más apropiado para comparar modelos con diferente número de variables independientes, es 45.2792%.

Para determinar si el modelo puede ser simplificado, note que el valor-P más alto para las pruebas de verosimilitud es 0.0008, que pertenece a ANGULO_Cat. Como el valor-P es menor que 0.05, ese término es estadísticamente significativo al nivel de confianza del 95.0%. Consecuentemente, usted probablemente no quisiera eliminar ninguna variable del modelo.

A continuación se presentan los resultados obtenidos del modelo base (Modelo 8) con la variable Número de Paraderos.

Tabla A.6.20 Modelo inicial con variable N° Paraderos (Máxima Verosimilitud)

Parámetro	Estimado	Error	
		Estandar	Razón de Momios Estimada
CONSTANTE	0.798662	0.140464	
N° PISTAS RS	0.128967	0.0272113	1.13765
ANGULO_Cat	0.00627225	0.00132545	1.00629
N° PARADEROS	0.0597084	0.0255984	1.06153
TMD_INTER_DIA	0.0000217729	0.00000202251	1.00002
TMD_CAM_DIA	0.0000369464	0.0000157992	1.00004
TMD_TXB_DIA	0.0000323768	0.0000125692	1.00003
DENSIDAD PEATONAL=1	-0.389486	0.0550692	0.677405
DENSIDAD PEATONAL=2	-0.22463	0.0560328	0.798812

Tabla A.6.21 Análisis de Desviación

Fuente	Desviación	Gl	Valor-P
Modelo	604.672	8	0.0000
Residuo	701.07	249	0.0000
Total (corr.)	1305.74	257	

Porcentaje de desviación explicado por el modelo = 46.3087

Porcentaje ajustado = 44.9302

Tabla A.6.22 Pruebas de Razón de Verosimilitud

Factor	Chi-Cuadrada	Gl	Valor-P
N° PISTAS RS	22.126	1	0.0000
ANGULO_Cat	24.3774	1	0.0000
N° PARADEROS	5.44055	1	0.0197
TMD_INTER_DIA	117.637	1	0.0000
TMD_CAM_DIA	5.15968	1	0.0231
TMD_TXB_DIA	6.57433	1	0.0103
DENSIDAD PEATONAL	49.7265	2	0.0000

Tabla A.6.23 Análisis de Residuos

	Estimación	Validación
n	258	
CME	307.405	
MAE	4.4062	
MAPE		
ME	0.0325893	

La salida muestra los resultados de ajustar un modelo de regresión de Poisson para describir la relación entre ACC_Diurno y 7 variable(s) independiente(s). La ecuación del modelo ajustado es:

$$\text{ACC_Diurno} = \exp(0.798662 + 0.128967 \text{ N}^\circ \text{ PISTAS RS} + 0.00627225 \text{ ANGULO_Cat} + 0.0597084 \text{ N}^\circ \text{ PARADEROS} + 0.0000217729 \text{ TMD_INTER_DIA} + 0.0000369464 \text{ TMD_CAM_DIA} + 0.0000323768 \text{ TMD_TXB_DIA} - 0.389486 \text{ DENSIDAD PEATONAL}=1 - 0.22463 \text{ DENSIDAD PEATONAL}=2)$$

Como el valor-P de la tabla de Análisis de Desviaciones es menor que 0.05, existe una relación estadísticamente significativa entre las variables, con un nivel de confianza del 95.0%. Además, el valor-P para los residuos es menor que 0.05, indicando que el modelo es significativamente peor que el mejor modelo posible para estos datos con un nivel de confianza del 95.0%.

Esta ventana también muestra que el porcentaje de desviación de ACC_Diurno explicado por el modelo es igual a 46.3087%. Este estadístico es similar al estadístico R-Cuadrada habitual. El porcentaje ajustado, que es más apropiado para comparar modelos con diferente número de variables independientes, es 44.9302%.

Para determinar si el modelo puede ser simplificado, note que el valor-P más alto para las pruebas de verosimilitud es 0.0231, que pertenece a TMD_CAM_DIA. Como el valor-P es menor que 0.05, ese término es estadísticamente significativo al nivel de confianza del 95.0%. Consecuentemente, usted probablemente no quisiera eliminar ninguna variable del modelo.

Al ver los resultados del modelo base con la variable Número de Paraderos, se decidió cambiar y probar entonces la modelación utilizando la variable Visibilidad. A continuación se presentan los resultados de éste modelo.

Tabla A.6.24 Modelo inicial con variable Visibilidad (Máxima Verosimilitud)

Parámetro	Estimado	Error	Razón de Momios
		Estandar	Estimada
CONSTANTE	0.976463	0.152577	
N° PISTAS RS	0.150158	0.0269055	1.16202
ANGULO_Cat	0.00515997	0.00138697	1.00517
VISIBILIDAD_Cat	-0.0418009	0.0159138	0.959061
TMD_INTER_DIA	0.0000219111	0.00000201914	1.00002
TMD_CAM_DIA	0.0000433933	0.000015984	1.00004
TMD_TXB_DIA	0.0000390508	0.0000122899	1.00004
DENSIDAD PEATONAL=1	-0.409053	0.0536934	0.664279
DENSIDAD PEATONAL=2	-0.249334	0.0549163	0.77932

Tabla A.6.25 Análisis de Desviación

Fuente	Desviación	Gl	Valor-P
Modelo	606.311	8	0.0000
Residuo	699.431	249	0.0000
Total (corr.)	1305.74	257	

Porcentaje de desviación explicado por el modelo = 46.4342

Porcentaje ajustado = 45.0557

Tabla A.6.26 Pruebas de Razón de Verosimilitud

Factor	Chi-Cuadrada	Gl	Valor-P
N° PISTAS RS	30.702	1	0.0000
ANGULO_Cat	14.6685	1	0.0001
VISIBILIDAD_Cat	7.07921	1	0.0078
TMD_INTER_DIA	119.787	1	0.0000
TMD_CAM_DIA	6.85552	1	0.0088
TMD_TXB_DIA	9.99075	1	0.0016
DENSIDAD PEATONAL	56.4617	2	0.0000

Tabla A.6.27 Análisis de Residuos

	Estimación	Validación
n	258	
CME	316.679	
MAE	4.39674	
MAPE		
ME	0.00837962	

La salida muestra los resultados de ajustar un modelo de regresión de Poisson para describir la relación entre ACC_Diurno y 7 variable(s) independiente(s). La ecuación del modelo ajustado es

$$\text{ACC_Diurno} = \exp(0.976463 + 0.150158 \text{ N}^\circ \text{ PISTAS RS} + 0.00515997 \text{ ANGULO_Cat} - 0.0418009 \text{ VISIBILIDAD_Cat} + 0.0000219111 \text{ TMD_INTER_DIA} + 0.0000433933 \text{ TMD_CAM_DIA} + 0.0000390508 \text{ TMD_TXB_DIA} - 0.409053 \text{ DENSIDAD PEATONAL=1} - 0.249334 \text{ DENSIDAD PEATONAL=2})$$

Como el valor-P de la tabla de Análisis de Desviaciones es menor que 0.05, existe una relación estadísticamente significativa entre las variables, con un nivel de confianza del 95.0%. Además, el valor-P para los residuos es menor que 0.05, indicando que el modelo es significativamente peor que el mejor modelo posible para estos datos con un nivel de confianza del 95.0%.

Esta ventana también muestra que el porcentaje de desviación de ACC_Diurno explicado por el modelo es igual a 46.4342%. Este estadístico es similar al estadístico R-Cuadrada habitual. El porcentaje ajustado, que es más apropiado para comparar modelos con diferente número de variables independientes, es 45.0557%.

Para determinar si el modelo puede ser simplificado, note que el valor-P más alto para las pruebas de verosimilitud es 0.0088, que pertenece a TMD_CAM_DIA. Como el valor-P es menor que 0.05, ese término es estadísticamente significativo al nivel de confianza del 95.0%. Consecuentemente, usted probablemente no quisiera eliminar ninguna variable del modelo.

Utilizando como base el modelo anterior se presentan los resultados del mismo, pero simplificado eliminando la variable menos significativa (TMD_CAM).

Tabla A.6.28 Modelo inicial con variable Visibilidad sin Flujo Camiones (Máxima Verosimilitud)

		Error	Razón de Momios
Parámetro	Estimado	Estandar	Estimada
CONSTANTE	0.950469	0.151865	
N° PISTAS RS	0.149756	0.026805	1.16155
ANGULO_Cat	0.00508678	0.00138457	1.0051
VISIBILIDAD_Cat	-0.039644	0.0158693	0.961132
TMD_INTER_DIA	0.0000254126	0.00000153387	1.00003
TMD_TXB_DIA	0.0000300625	0.0000117815	1.00003
DENSIDAD PEATONAL=1	-0.401952	0.0535863	0.669013
DENSIDAD PEATONAL=2	-0.249635	0.0549101	0.779085

Tabla A.6.29 Análisis de Desviación

Fuente	Desviación	Gl	Valor-P
Modelo	599.455	7	0.0000
Residuo	706.286	250	0.0000
Total (corr.)	1305.74	257	

Porcentaje de desviación explicado por el modelo = 45.9092

Porcentaje ajustado = 44.6838

Tabla A.6.30 Pruebas de Razón de Verosimilitud

Factor	Chi-Cuadrada	Gl	Valor-P
Nº PISTAS RS	30.7544	1	0.0000
ANGULO_Cat	14.2842	1	0.0002
VISIBILIDAD_Cat	6.39347	1	0.0115
TMD_INTER_DIA	261.0	1	0.0000
TMD_TXB_DIA	6.44365	1	0.0111
DENSIDAD PEATONAL	54.6072	2	0.0000

Tabla A.6.31 Análisis de Residuos

	Estimación	Validación
n	258	
CME	321.999	
MAE	4.50124	
MAPE		
ME	0.00658133	

La salida muestra los resultados de ajustar un modelo de regresión de Poisson para describir la relación entre ACC_Diurno y 6 variable(s) independiente(s). La ecuación del modelo ajustado es:

$$\text{ACC_Diurno} = \exp(0.950469 + 0.149756 \text{ N}^\circ \text{ PISTAS RS} + 0.00508678 \text{ ANGULO_Cat} - 0.039644 \text{ VISIBILIDAD_Cat} + 0.0000254126 \text{ TMD_INTER_DIA} + 0.0000300625 \text{ TMD_TXB_DIA} - 0.401952 \text{ DENSIDAD PEATONAL}=1 - 0.249635 \text{ DENSIDAD PEATONAL}=2)$$

Como el valor-P de la tabla de Análisis de Desviaciones es menor que 0.05, existe una relación estadísticamente significativa entre las variables, con un nivel de confianza del 95.0%. Además, el valor-P para los residuos es menor que 0.05, indicando que el modelo es significativamente peor que el mejor modelo posible para estos datos con un nivel de confianza del 95.0%.

Esta ventana también muestra que el porcentaje de desviación de ACC_Diurno explicado por el modelo es igual a 45.9092%. Este estadístico es similar al estadístico R-Cuadrada habitual. El porcentaje ajustado, que es más apropiado para comparar modelos con diferente número de variables independientes, es 44.6838%.

Para determinar si el modelo puede ser simplificado, note que el valor-P más alto para las pruebas de verosimilitud es 0.0115, que pertenece a VISIBILIDAD_Cat. Como el valor-P es menor que 0.05, ese término es estadísticamente significativo al nivel de confianza del 95.0%. Consecuentemente, usted probablemente no quisiera eliminar ninguna variable del modelo.



