

APLICAÇÃO DE MODELOS DE PREVISÃO DE ACIDENTES RODOVIÁRIOS À REDE ASCENDI

JOANA CABRAL CASTELO BRANCO RIBEIRO

Dissertação submetida para satisfação parcial dos requisitos do grau de
MESTRE EM ENGENHARIA CIVIL — ESPECIALIZAÇÃO EM VIAS DE COMUNICAÇÃO

Orientadora: Professora Doutora Sara Maria Pinho Ferreira

JUNHO DE 2021

MESTRADO INTEGRADO EM ENGENHARIA CIVIL 2020/2021

DEPARTAMENTO DE ENGENHARIA CIVIL

Tel. +351-22-508 1901

Fax +351-22-508 1446

✉ miec@fe.up.pt

Editado por

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Rua Dr. Roberto Frias

4200-465 PORTO

Portugal

Tel. +351-22-508 1400

Fax +351-22-508 1440

✉ feup@fe.up.pt

🌐 <http://www.fe.up.pt>

Reproduções parciais deste documento serão autorizadas na condição que seja mencionado o Autor e feita referência a *Mestrado Integrado em Engenharia Civil - 2020/2021 - Departamento de Engenharia Civil, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal, 2021.*

As opiniões e informações incluídas neste documento representam unicamente o ponto de vista do respetivo Autor, não podendo o Editor aceitar qualquer responsabilidade legal ou outra em relação a erros ou omissões que possam existir.

Este documento foi produzido a partir de versão eletrónica fornecida pelo respetivo Autor.

A meus Pais

O conhecimento une cada um consigo mesmo e todos com todos.

José Saramago

AGRADECIMENTOS

À minha mãe, Maria de Lourdes, o pilar principal na minha vida, que sempre me apoiou em todos os momentos e me incentivou a fazer mais e melhor.

Ao meu irmão, Diogo, que está sempre ao meu lado e é o meu exemplo.

Ao meu pai, José Mário, por apoiar as minhas escolhas ao longo desta jornada.

Um agradecimento especial, à minha orientadora, Professora Doutora Sara Ferreira, por me ter acompanhado ao longo destes seis meses e por ter estado sempre presente no desenvolvimento de todas as etapas deste trabalho. Estou grata por toda a sua ajuda e dedicação.

Ao engenheiro João Neves, pela sua coorientação, agradeço a oportunidade de adquirir conhecimento na área da segurança rodoviária e os pelos seus ensinamentos transmitidos ao longo do estágio. Agradeço a confiança que depositou em mim e no meu trabalho.

A todos os profissionais da empresa Ascendi, por todo o apoio que me deram para ultrapassar as dificuldades que foram surgindo ao longo deste trabalho. Em especial, ao engenheiro Carlos Oliveira e à engenheira Telma Silva, do serviço de segurança rodoviária, por me terem acolhido e acompanhado ao longo deste estágio, ajudando em todos os momentos necessários. Queria agradecer aos meus colegas, Tiago, Samuel e Suellen, pelos momentos partilhados e pela ajuda constante. Desejo tudo de bom para todos e muito sucesso.

Aos meus amigos da vida, aos que me acompanham desde pequena e aos que tive o privilégio de conhecer ao longo da vida. A todos eles agradeço por estarem sempre ao meu lado. Agradeço à namorada do meu irmão e amiga, Maria Pontes, por me incentivar a ser melhor e motivar-me todos os dias. Aos meus companheiros da faculdade, Afonso, Linhares e Diogo, agradeço por partilharem a melhor fase comigo.

Por último, obrigada a todas as pessoas que me ajudaram direta ou indiretamente nesta jornada.

RESUMO

A modelação de frequência de acidentes é uma ferramenta importante para a gestão da segurança rodoviária. O aumento da procura dos meios de transporte, em especial, do transporte rodoviário, tem resultado num claro aumento do número de acidentes nessa envolvente. Naturalmente, um dos grandes desafios para os investigadores e gestores rodoviários passa pela redução dos acidentes nas estradas. Nesse sentido, esta dissertação apresenta um estudo de modelos de previsão de acidentes e teve em consideração modelos lineares generalizados com distribuição probabilística de Poisson, dado que os acidentes são acontecimentos aleatórios. Este modelo tem como objetivo apoiar a avaliação da segurança de locais, identificar e classificar locais perigosos ou zonas de acumulação de acidentes, bem como, estudar o efeito de medidas aplicadas na melhoria da segurança. Num contexto prático, este trabalho tem como foco a construção e análise de um modelo de previsão de frequência de acidentes, referente à autoestrada A44, em Portugal, concessionada à empresa Ascendi. Os modelos de previsão foram desenvolvidos para duas variáveis dependentes distintas, acidentes totais e acidentes com vítimas, para um período de análise de seis anos (2015 a 2020). Para fins de validação e comparação de dois anos distintos, no que toca à mobilidade e transportes, foi realizado um segundo modelo de previsão, que foi desenvolvido para um período de quatro anos (2015 a 2018), permitindo a validação e análise comparativa dos anos 2019 e 2020. Considerou-se, ainda, a comparação destes resultados com o caso da autoestrada A29, estudada anteriormente (Souza, S. 2021). A análise comparativa foi positiva e revelou uma estabilidade dos modelos, mesmo quando aplicados num ano atípico.

PALAVRAS-CHAVE: Modelo de previsão de acidentes; Acidentes com vítimas; Autoestradas; Segurança rodoviária; Validação .

ABSTRACT

Accident frequency modeling is an important tool for road safety management. The increase in demand for means of transport, especially road transport, has resulted in a significant increase in the number of accidents in this environment. Hence, one of the great challenges for researchers and road managers, is to reduce accidents on the road. Thus, this dissertation presents a study of accident prediction models and took into consideration generalized linear models with Poisson's probabilistic distribution, since accidents are random events. This model aims to support the evaluation of site safety, the identification and classification of dangerous sites or accident accumulation zones and study the effect of measures applied to improve safety. In a practical context, this work focuses on the construction and analysis of an accident frequency prediction model, regarding the A44 highway in Portugal, under concession to the company Ascendi. The prediction models were developed for two distinct dependent variables, total accidents and accidents with casualties, for a six-year analysis period (2015 to 2020). For validation purposes and comparison of two different years, as far as mobility and transportation is concerned, a second prediction model, which was developed for a four-year period (2015-2018), was carried out, allowing the validation and comparative analysis of the years 2019 and 2020. The comparison of these results with the case of highway A29, analyzed in a previous study, has also been considered (Souza, S. 2021). The comparative analysis was positive and proved the stability of the models, even when applied in an atypical year.

KEYWORDS: Accident prediction models; Casualty accidents; Highways; Road safety; Validation.

ÍNDICE GERAL

AGRADECIMENTOS.....	
RESUMO.....	ii
ABSTRACT.....	iv
1. INTRODUÇÃO	1
1.1. Preâmbulo	1
1.2. Objetivos	5
1.3. Estrutura da dissertação.....	6
2. MODELOS DE PREVISÃO DE ACIDENTES.....	7
2.1. Enquadramento	7
2.2. Características dos Acidentes.....	8
2.3. Modelos de Frequência de Acidentes	10
2.3.1. Modelos de Regressão de Poisson e Binomial Negativo.....	11
2.3.2. Segmentação da Rede.....	18
2.4. Validação.....	19
2.4.1. Medidas de Desempenho.....	20
2.4.2. Análise Gráfica.....	23
3. CASO DE ESTUDO	25
3.1. Metodologia.....	25
3.2. Enquadramento do caso de estudo.....	26
3.2.1. Definição da Área de Estudo.....	30
3.2.2. Análise da sinistralidade da A44.....	31

3.2.2. Recolha dos Dados.....	34
3.3. Segmentação das vias.....	36
3.4. Dados considerados na modelação	38
3.4.1. Correlação entre Variáveis	42
3.5. Modelação	43
3.6. Validação do Modelo	44
4. ANÁLISE DE RESULTADOS	45
4.1. Modelo de Previsão de Frequência de Acidentes totais – A44 (15-20).....	46
4.1.1. Etapa de Modelação	46
4.1.2. Etapa de Avaliação do Desempenho do Modelo	50
4.2. Modelo de Previsão de Frequência de Acidentes com Vítimas – A44 (15-20)..	52
4.2.1. Etapa de Modelação	52
4.2.2. Etapa de Avaliação do Desempenho do Modelo	54
4.3. Validação dos Modelos de Previsão de Acidentes.....	56
4.3.1. Estimação do Modelo de Previsão de Frequência de Acidentes	57
Autoestrada A44	57
Autoestrada A29	59
4.3.2. Validação dos modelos de previsão de frequência de acidentes	59
5. CONCLUSÕES	77
Referências Bibliográficas.....	81
ANEXO A - Tabela matriz de correlação.....	87
ANEXO B - Resultados dos modelos <i>Software</i> nlogit versão 5	89

Anexo B.1 – Modelo de frequência de acidentes totais - A44 (15-20)	90
Resultado Poisson	90
Resultado Binomial Negativo	91
Anexo B.2 – Modelo de frequência de acidentes com vítimas – A44 (15-20).....	92
Resultado Poisson	92
Resultado Binomial Negativo	93
Anexo B.3 – Modelo de frequência de acidentes totais– A44 (15-18)	94
Anexo B.4 – Modelo de frequência de acidentes com vítimas – A44 (15-18).....	95
Anexo B.5 – Modelo de frequência de acidentes totais– A29 (15-17)	96

ÍNDICE DE FIGURAS

Figura 1.1: Número de acidentes mortais por ano na UE desde 2000 (Cinzento: Observado, Azul: Objetivo). Fonte: CARE (Base de dados de acidentes rodoviários EU).....	2
Figura 1.2: Evolução dos AcV e AcVM+AcFG no continente, entre 2016 e 2020. Fonte: ANSR, 2020.....	4
Figura 1.3: Evolução das vítimas por tipo de lesão no continente, entre 2016 e 2020. Fonte: ANSR, 2020.....	4
Figura 1.4: Evolução dos Acidentes (Ac) e Acidentes com vítimas (AcV) no continente, entre 2016 e 2020.....	5
Figura 1.5: Evolução das vítimas por tipo de lesão, entre 2016 e 2020 (FL: ferido ligeiro; FG: ferido grave; VM: vítima mortal).....	5
Figura 2.1: Regressão à média no “antes” e “depois”. Fonte: Manual de Segurança Rodoviária (HSM).....	9
Figura 3.1: Estrutura da metodologia.....	26
Figura 3.2: Concessões Ascendi. Fonte: Anuário de Segurança Rodoviária Ascendi, 2020.....	27
Figura 3.3: Percentagens de Acidentes Totais por Concessão [2015-2020].....	28
Figura 3.4: Percentagens de Acidentes com Vítimas por Concessão [2015-2020].....	29
Figura 3.5: Gráfico de acidentes totais, acidentes com vítimas e relação sinistralidade/extensão, nas autoestradas da CP [2015-2020].....	29
Figura 3.6: Mapa da Concessão Costa de Prata, AE A44. Fonte: SIG Ascendi.....	31
Figura 3.7: Gráfico de sinistralidade na A44 [2015 – 2020].....	32
Figura 3.8: Gráfico do TMDA na A44 [2015 – 2020].....	32
Figura 3.9: Localização dos acidentes rodoviários (Verde: acidentes s/ vítimas, Vermelho: acidentes c/ vítimas). Fonte: SIG Ascendi.....	33
Figura 3.10: Tráfego Médio Diário Anual por sublanço e sentido [2015 – 2020]. Fonte: Ascendi.....	34
Figura 3.11: Gráfico do Tráfego Médio Diário Anual na A44 [2015 – 2020].....	35
Figura 4.1: Gráfico resíduos acumulados da variável TMDA para acidentes totais.....	51

Figura 4.2: Gráfico resíduos acumulados da variável comprimento para acidentes totais.....	51
Figura 4.3: Gráfico resíduos acumulados da variável TMDA para acidentes com vítimas.....	55
Figura 4.4: Gráfico resíduos acumulados da variável comprimento para acidentes com vítimas.....	55
Figura 4.5: Gráfico resíduos acumulados da variável TMDA para acidentes totais na A44 [2019].....	62
Figura 4.6: Gráfico resíduos acumulados da variável comprimento para acidentes totais na A44 [2019].....	63
Figura 4.7: Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A44 [2019].....	64
Figura 4.8: Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A44 [2019].....	64
Figura 4.9: Gráfico resíduos acumulados da variável TMDA para acidentes totais na A44 [2020].....	65
Figura 4.10: Gráfico resíduos acumulados da variável comprimento para acidentes totais na A44 [2020].....	65
Figura 4.11: Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A44 [2020].....	66
Figura 4.12: Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A44 [2020].....	67
Figura 4.13: Gráfico resíduos acumulados da variável TMDA para acidentes com vítimas na A44 [2019].....	67
Figura 4.14: Gráfico resíduos acumulados da variável comprimento para acidentes com vítimas na A44 [2019].....	68
Figura 4.15: Gráfico dispersão dos valores para a variável TMDA para acidentes com vítimas na A44 [2019].....	69
Figura 4.16: Gráfico dispersão dos valores para a variável comprimento para acidentes com vítimas na A44 [2019].....	69
Figura 4.17: Gráfico resíduos acumulados da variável TMDA para acidentes com vítimas na A44 [2020].....	70

Figura 4.18: Gráfico resíduos acumulados da variável comprimento para acidentes com vítimas na A44 [2020].....	70
Figura 4.19: Gráfico dispersão dos valores para a variável TMDA para acidentes com vítimas na A44 [2020].....	71
Figura 4.20: - Gráfico dispersão dos valores para a variável comprimento para acidentes com vítimas na A44 [2020].....	71
Figura 4.21: - Gráfico resíduos acumulados da variável TMDA para acidentes totais na A29 [2019].....	72
Figura 4.22: - Gráfico resíduos acumulados da variável comprimento para acidentes totais na A29 [2019].....	72
Figura 4.23: - Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A29 [2019].....	73
Figura 4.24: - Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A29 [2019].....	74
Figura 4.25: - Gráfico resíduos acumulados da variável TMDA para acidentes totais na A29 [2020].....	74
Figura 4.26: - Gráfico resíduos acumulados da variável comprimento para acidentes totais na A29 [2020].....	75
Figura 4.27: - Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A29 [2020].....	76
Figura 4.28: - Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A29 [2020].....	76

ÍNDICE DE QUADROS

Quadro 1.1: N° de acidentes e vítimas ocorridos no continente e regiões autónomas durante 2020. Fonte: ANSR, 2020.....	3
Quadro 1.2: Evolução da sinistralidade no continente, entre 2016 e 2020. Fonte: ANSR, 2020....	3
Quadro 3.1: Concessões Ascendi.....	28
Quadro 3.2: Sublanços da A44.....	35
Quadro 3.3: Descrição das variáveis.....	40
Quadro 3.4: Legenda de categorias.....	41
Quadro 3.5: Descrição estatísticas das variáveis independentes.....	41
Quadro 4.1: Resultados das estimativas do modelo Poisson para acidentes totais.....	47
Quadro 4.2: Elasticidade das variáveis contínuas - Acidentes Totais.....	48
Quadro 4.3: Pseudo-elasticidade das variáveis dummies - Acidentes Totais.....	49
Quadro 4.4: Resultados das estimativas do modelo Poisson para acidentes com vítimas.....	52
Quadro 4.5: Elasticidade das variáveis contínuas - Acidentes com Vítimas.....	53
Quadro 4.6: Pseudo-elasticidade das variáveis dummies - Acidentes com Vítimas.....	53
Quadro 4.7: Resultados das estimativas do modelo Poisson para acidentes totais – A44.....	57
Quadro 4.8: Resultados das estimativas do modelo Poisson para acidentes com vítimas – A44....	58
Quadro 4.9: Resultados das estimativas do modelo Poisson para acidentes totais – A29.....	59
Quadro 4.10: Erros relativos e coeficientes de determinação modelos de previsão [2019 – 2020].....	60
Quadro 4.11: Estatísticas resumidas dos valores observados e previstos [2019 – 2020].....	60
Quadro 4.12: Medidas de desempenho modelos de previsão [2019 - 2020].....	61

SÍMBOLOS, ACRÓNIMOS E ABREVIATURAS

ANSR - Autoridade Nacional de Segurança Rodoviária

CARE - Base de dados comunitária sobre Acidentes nas Estradas na Europa

CE - Comissão Europeia

DEC - Departamento de Engenharia Civil

ENSR - Estratégia Nacional de Segurança Rodoviária

HSM – Manual de Segurança Rodoviária

IFT - Fórum Internacional de Transportes

MPA – Modelo de Previsão de Acidentes

OCDE – Organização para a Cooperação e Desenvolvimento Económico

SIG – Sistema de Informação Geográfica

SS – Sistema Seguro

SR – Segurança Rodoviária

TMDA - Tráfego Médio Diário Anual

UE – União Europeia

ZAA - Zona de Acumulação de Acidentes

1

INTRODUÇÃO

1.1. PREÂMBULO

Atualmente, a sinistralidade rodoviária é um dos problemas mais graves, sendo considerada uma questão de saúde pública. Os acidentes rodoviários são a principal causa de morte a nível mundial na faixa etária mais jovem, e a oitava para todas as idades, apresentando uma carga social e económica nas comunidades (ANSR, 2020).

Segundo a Comissão Europeia (CE), morre em acidentes rodoviários, por ano, 1,25 milhões de pessoas em todo o mundo, das quais 26000 na União Europeia (UE). A UE é considerada um caso de sucesso, devido a muitos fatores, dos quais: estabelecimento da segurança rodoviária (SR) como prioridade política; a participação de todos os países da comunidade para manter um espaço seguro; adoção da abordagem “sistema seguro” e da perspetiva “visão zero”; fixação de objetivos, recolha de dados e monitorização continua dos resultados; assim como ações dedicadas a campanhas de educação e sensibilização.

A abordagem de Sistema Seguro (SS) tem-se destacado na problemática da segurança rodoviária. Esta integra os principais elementos de gestão e áreas de ação, com o objetivo de criar um sistema de mobilidade seguro. Segundo a OCDE e IFT (2008), este é o único caminho para alcançar a “Visão Zero”. A “Visão Zero” é fundamentada por uma base ética, apoiado numa nova perspetiva relativamente ao problema da segurança rodoviária e respetiva resolução.

Assim, este é o caminho ambicioso da UE, a longo prazo, para alcançar zero mortes ou lesões graves, provocadas pelos acidentes rodoviários, até 2050.

Nos últimos anos tem-se observado que a progressão da redução das taxas de mortalidade rodoviária na UE estagnou – figura 1.1. Segundo a CE, no ano 2020, observou-se uma diminuição do número de mortes nas estradas. Estima-se que em 2020, 18800 pessoas morreram nas estradas, resultando num decréscimo de 17%, relativamente a 2019. Isto significa uma diminuição de quase 4000 mortes rodoviárias, na UE em 2020, comparativamente a 2019. No contexto das medidas tomadas no combate à pandemia Covid-19, o volume de tráfego foi menor, apresentando um impacto claro no número de mortes nas estradas.

Contudo, Adina Vălean afirma que “apesar das estradas da UE continuarem a ser as mais seguras do mundo, apresentam-se atrás do objetivo estipulado para a última década, sendo necessária uma ação conjunta para evitar regressar aos níveis pré-covid”. (CE, 2021)

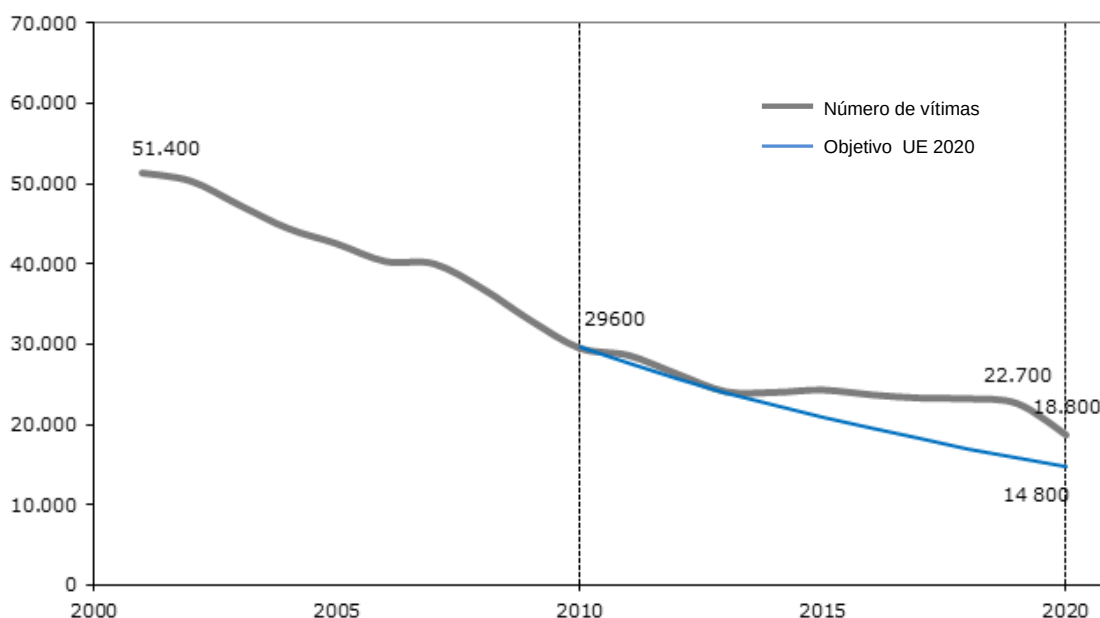


Figura 1.1: Número de acidentes mortais por ano na UE desde 2000 (Cinzento: Observado, Azul: Objetivo). Fonte: CARE (Base de dados de acidentes rodoviários EU)

A segurança dos utilizadores do sistema rodoviário (condutores e passageiros) constitui, cada vez mais, um dos principais objetivos no planeamento e operação deste meio de transporte (Sinay & Tamayo, 2005). Deste modo, as empresas e instituições que atuam na segurança rodoviária, apresentam uma grande responsabilidade na implementação de medidas que incluem, entre outras, medidas que visam melhorar a conceção e operação das vias. Estas medidas podem apresentar uma importante função na redução da frequência de acidentes, assim como, na gravidade dos ferimentos sofridos pelos seres humanos envolvidos.

Em Portugal, a Autoridade Nacional de Segurança Rodoviária (ANSR) é a instituição responsável pela gestão da segurança rodoviária. Esta entidade tem como missão o planeamento e coordenação no território português, no que diz respeito à segurança rodoviária, bem como a aplicação do direito contraordenacional rodoviário.

Segundo o relatório anual de sinistralidade, da ANSR, de 2020, a sinistralidade rodoviária anual foi condicionada, assim como outros fenómenos da sociedade, tanto a nível nacional, como a nível mundial. As medidas de confinamento tomadas para conter a pandemia do SARS Cov-2 alteraram, de forma global, a mobilidade e, naturalmente, a exposição ao risco de acidente, assim como, o comportamento dos utentes das vias.

Deste modo, quando comparada a sinistralidade de 2019 com 2020, e de acordo com a quadro 1.1, observou-se uma melhoria nos principais indicadores de sinistralidade no Continente: menos 9.203 acidentes (-25,8%), menos 84 vítimas mortais (-17,7%), menos 472 feridos graves (-20,5%) e menos 12.496 feridos leves (-28,9%).

Quadro 1.1: Nº de acidentes e vítimas ocorridos no continente e regiões autónomas durante 2020. Fonte: ANSR, 2020

Sinistralidade, 2020 vs 2019												
	AcV			VM			FG			FL		
	2019	2020	$\Delta(\%)$ 20/19	2019	2020	$\Delta(\%)$ 20/19	2019	2020	$\Delta(\%)$ 20/19	2019	2020	$\Delta(\%)$ 20/19
Continente	35704	26501	-25,8%	474	390	-17,7%	2301	1829	-20,5%	43202	30706	-28,9%
Açores	611	501	-18,0%	7	6	-14,3%	120	100	-16,7%	686	538	-21,6%
Madeira	936	723	-22,8%	39	8	-79,5%	111	67	-39,6%	1065	827	-22,3%
Total	37251	27725	-25,6%	520	404	-22,3%	2532	1996	-21,2%	44953	32071	-28,7%

Nas figuras 1.2 e 1.3 pode-se observar a evolução da sinistralidade, através dos seus indicadores, para os anos de 2016 a 2020. Realizando uma análise focado no Continente e comparando o ano 2019 e 2020, é possível observar que todos os indicadores apresentaram melhores resultados no ano 2020: -18,0% nos acidentes, -10,3% nos acidentes com mortos e/ou feridos graves, -12,4% em termos de vítimas mortais, -13,0% de feridos graves e -21,5% de feridos leves. Contudo, o índice de gravidade aumentou 6,8% de 1,38 para 1,47 (quadro 1.2).

Quadro 1.2: Evolução da sinistralidade no continente, entre 2016 e 2020. Fonte: ANSR, 2020

Ano	AcV	AcVM+AcFG	AcVM	VM	FG	FL	IG
2016	32299	2201	416	445	2102	39121	1,38
2017	34416	2397	488	510	2198	41787	1,48
2018	34235	2337	468	508	2141	41356	1,48
2019	35704	2403	429	474	2301	43202	1,33
2020	26501	1975	372	390	1829	30706	1,47

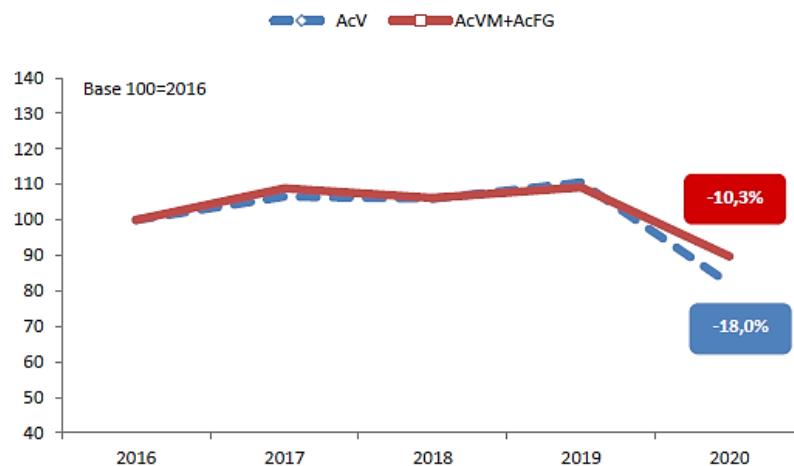


Figura 1.2: Evolução dos AcV e AcVM+AcFG no continente, entre 2016 e 2020. Fonte: ANSR, 2020

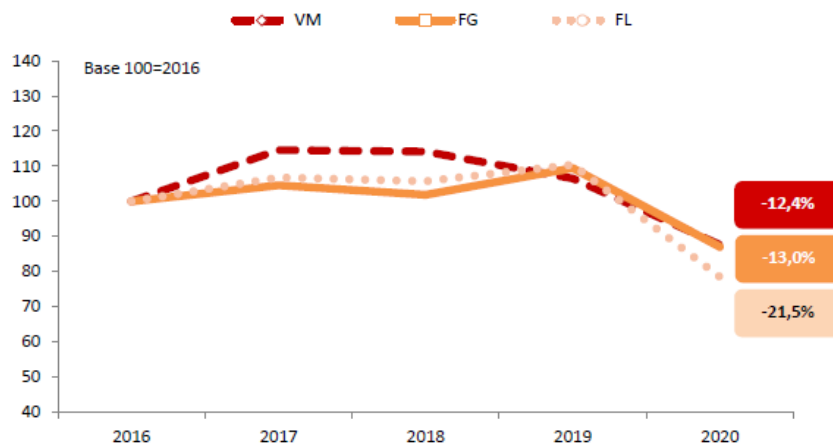


Figura 1.3: Evolução das vítimas por tipo de lesão no continente, entre 2016 e 2020. Fonte: ANSR, 2020

As entidades responsáveis pela gestão da via têm um papel fundamental na diminuição da sinistralidade rodoviária. É título de exemplo a Ascendi, na qual o presente trabalho foi realizado, que é responsável pela operação e manutenção de infraestruturas, atua como gestora de ativos, e é prestadora de serviços de cobrança de portagens.

Nas figuras 1.4 e 1.5 pode-se observar a evolução da sinistralidade na rede Ascendi, analisando os indicadores para os anos de 2016 a 2020. Através da análise gráfica é possível comparar o ano de 2019 e 2020 e comprovar uma diminuição de 27% nos acidentes e 33% nos acidentes com vítimas. No que diz respeito às vítimas por tipo de lesão, notou-se um decréscimo de 38% nos feridos ligeiros, 50% nos feridos graves e 7% nas vítimas mortais.

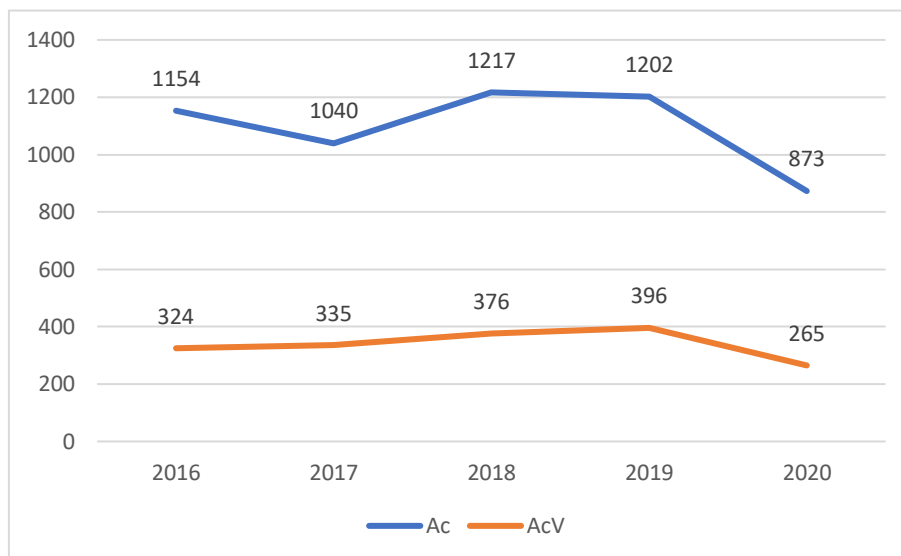


Figura 1.4: Evolução dos Acidentes (Ac) e Acidentes com vítimas (AcV) no continente, entre 2016 e 2020

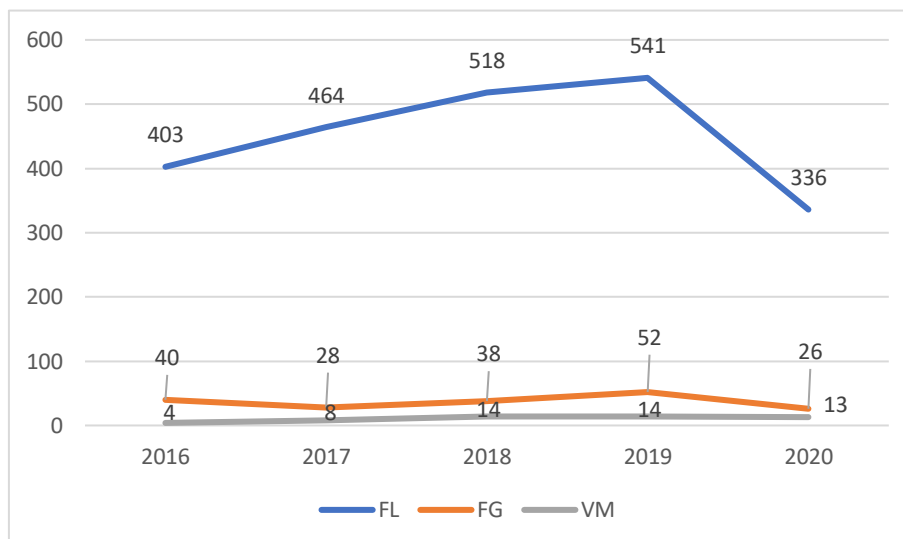


Figura 1.5: Evolução das vítimas por tipo de lesão, entre 2016 e 2020 (FL: ferido ligeiro; FG: ferido grave; VM: vítima mortal)

1.2. OBJETIVOS

Este trabalho tem como objetivo o desenvolvimento de modelos de previsão de frequência de acidentes, incluindo calibração e validação, para as autoestradas da rede viária da Ascendi. No âmbito da segurança rodoviária, este estudo permite, através de métodos estatísticos, relacionar os acidentes totais e com vítimas, com uma série de variáveis explicativas (fatores contribuintes), criando uma ferramenta robusta que permite por um lado, avaliar o efeito dessas variáveis na sinistralidade e por outro, estimar a sinistralidade em cenários específicos.

O estudo realizado neste trabalho insere-se na continuação do trabalho anterior, no âmbito da dissertação de Souza, S. (2021) intitulado por “Identificação e análise de fatores de risco em autoestrada: aplicação de modelos estatísticos”. No trabalho anterior, foram desenvolvidos modelos de previsão de acidentes para a autoestrada A29, pretendendo-se com o presente trabalho alargar o estudo à autoestrada A44. Ainda, de forma a aprofundar o processo de validação dos modelos, são aplicadas medidas de desempenho e análises gráficas, com o objetivo de validar os modelos desenvolvidos. A análise comparativa das autoestradas permitiu estudar a robustez dos modelos de previsão de acidentes, apesar das características heterogêneas das mesmas e a sua possível utilização como ferramenta eficaz na aplicação de medidas na segurança rodoviária.

No âmbito do processo de validação dos modelos, foi considerado o caso particular do ano 2020, caracterizado pela crise pandémica, com o objetivo de estudar se, para um ano disruptivo, os modelos matemáticos tradicionais são capazes de prever os acidentes de forma ajustada. Esta análise particular pretende contribuir para o estudo da aplicabilidade destas ferramentas, em contexto de crise, na área da mobilidade e segurança.

Com o objetivo de agilizar procedimentos e estender o desenvolvimento de modelos a toda a rede de autoestradas, geridas pela empresa Ascendi, os dados de toda a rede foram tratados, de acordo com procedimentos específicos, para que, no futuro, a aplicação de modelos seja sistematizada.

1.3. ESTRUTURA DA DISSERTAÇÃO

A estrutura da dissertação encontra-se dividida por 5 capítulos:

- Capítulo 1: introdução e breve exposição do tema e âmbito de estudo, objetivo e estrutura da dissertação;
- Capítulo 2: enquadramento sobre os acidentes, contextualização relativamente aos modelos de previsão de acidentes, suas características e processo de validação;
- Capítulo 3: enquadramento no caso de estudo, metodologia e definição do modelo;
- Capítulo 4: apresentação, análise e comparação dos resultados obtidos através de processos de validação;
- Capítulo 5: exibição das principais conclusões retiradas do estudo e recomendações futuras para possíveis melhorias.

2

MODELOS DE PREVISÃO DE ACIDENTES

2.1. ENQUADRAMENTO

Ao longo dos tempos tem vindo a ser objeto de análise o crescimento dos sistemas de transporte, em especial o rodoviário, conduzindo a um aumento dos acidentes rodoviários. A segurança dos utilizadores do sistema (condutores, passageiros e pedestres) constitui um dos objetivos principais do planeamento e operação deste sistema de transporte (Sinay & Tamayo, 2005).

Partindo do princípio fundamental de que os condutores cometerão sempre erros, as infraestruturas deverão contribuir para a diminuição desses erros e minimizar a gravidade das suas consequências. Sendo o âmbito deste trabalho as infraestruturas, é de destacar as primeiras medidas aplicadas a este nível, que tiveram como objetivo a identificação dos locais onde se registam o maior número de acidentes, de forma a corrigir os problemas detetados.

A par com os países europeus, Portugal aplica medidas de tratamento dos pontos negros, em simultâneo com a melhoria da qualidade das infraestruturas, nomeadamente, a nível da sinalização. Contudo, estas medidas, apesar de se terem revelado positivas, são geralmente aplicadas após a ocorrência de um número elevado de acidentes.

Em Portugal, a definição de ponto negro, segundo a ANSR, corresponde ao lanço de estrada com o máximo de 200 metros de extensão, no qual se registou, pelo menos, 5 acidentes com vítimas, no ano em análise, e cuja soma de indicadores de gravidade é superior a 20. Em alternativa a esta definição específica, existe a definição de zona de acumulação de acidentes (ZAA), que corresponde a uma zona geográfica onde, por influência de características da infraestrutura rodoviária específicas à área, a frequência esperada de acidentes é superior ao expectável face à

distribuição de acidentes nas áreas circundantes, nomeadamente em zonas aparentemente semelhantes (Cardoso, 1998).

Assim, o método de deteção de ZAA tem por base a comparação da frequência de acidentes num determinado local com a frequência expectável média do conjunto de elementos similares. Este método de identificação e posterior tratamento, aplicando medidas de engenharia, nas ZAAs, é utilizado em muitos países por toda a Europa. Sayed, T. and Leur, P. d. (2000) referem que, apesar do tratamento de ZAA ter-se revelado relevante e demonstrado ótimos resultados, é importante evoluir para uma abordagem proativa na SR.

Segundo Lord e Mannering (2010), tem-se investigado fatores que afetam o número de acidentes que ocorrem num espaço geográfico, normalmente numa interseção ou segmento de via, durante um período específico. Destes resultam dados de frequência de acidentes e/ou gravidade, que é um dos objetivos principais do desenvolvimento de modelos de previsão de acidentes.

2.2. CARACTERÍSTICAS DOS ACIDENTES

Os acidentes rodoviários são acontecimentos raros e aleatórios, que resultam em ferimentos e danos materiais. Estes correspondem a uma pequena percentagem dos fenómenos que ocorrem numa rede viária e são o resultado da convergência de vários eventos, influenciados por diferentes fatores (período do dia, velocidade, condições do veículo, estado do condutor, etc.) (AASHTO, 2010). Assim sendo, a análise dos acidentes tem em conta dois grupos de fatores:

- Fatores de exposição ao risco de acidentes (volume de tráfego e distância percorrida);
- Fatores associados aos condutores, veículos, via e ambiente viário.

Relativamente aos fatores que originam um acidente na rede viária, estes podem ser em parte controlados, assim como aleatórios e inesperados.

No que diz respeito aos fatores que podem ser controlados, como por exemplo, os relacionados com o condutor e veículo, e que são representados por diferentes propriedades (tempo de reação, eficácia na travagem, controlo da velocidade), estes podem demarcar a ocorrência ou não de um acidente. O potencial risco de ocorrência de acidente aumenta, particularmente, com a alteração inesperada do fluxo de tráfego numa autoestrada, um atravessamento de um peão numa estrada ou a existência de um objeto na mesma. Todavia, estes eventos críticos podem ser controlados, evitando o potencial risco de acidente, através de ações antecipadas por parte do condutor.

Não obstante, as circunstâncias que originam um acidente em determinado momento não são, necessariamente, as mesmas que causam outro evento similar. Esta observação reflete a aleatoriedade inerente aos acidentes.

Tendo em conta as características dos acidentes acima descritas, considera-se apropriada a aplicação de técnicas estatísticas para o estudo da segurança rodoviária, que consistem na elaboração de Modelos de Previsão de Acidentes (MPA), com base no conhecimento histórico dos acidentes ocorridos no local em estudo (Claude, 2012). Os modelos têm como objetivo determinar um valor esperado de acidentes, com determinadas características, para uma dada unidade de tempo e local. Estes integram modelos probabilísticos que permitem calcular a frequência com que esses eventos raros, sob determinada condição, acontecem (Cunto, 2009).

A frequência de acidentes é a base fundamental para a análise da segurança, seleção dos locais para tratamento e avaliação dos efeitos dos tratamentos. Contudo, é necessário conhecer as características dos dados dos acidentes, que podem afetar a fiabilidade da análise.

O estudo da segurança dos locais com elevado número de acidentes, “antes” e “depois”, da aplicação de medidas, é uma prática comum para determinar a eficácia das mesmas. Porém, essas avaliações podem ser enviesadas pelo fenómeno de regressão à média. Este enviesamento ocorre quando se seleciona um local para estudo, baseado na tendência de frequência de acidentes a curto prazo. Analisando a figura 2.1, podemos observar o efeito de regressão à média no estudo antes e depois. Pela figura, observa-se que um local com uma elevada frequência de acidentes, num determinado período, é suscetível de ter uma diminuição da frequência acidentes durante o período seguinte, mesmo que não sejam tomadas medidas (medidas que resultam numa redução de acidentes). Devido a este facto, é importante ter em consideração o efeito do fenómeno da regressão à média, para determinar a redução da frequência de acidentes que ocorreria sem a aplicação de medidas e para que não exista uma falsa percepção de eficácia da medida.

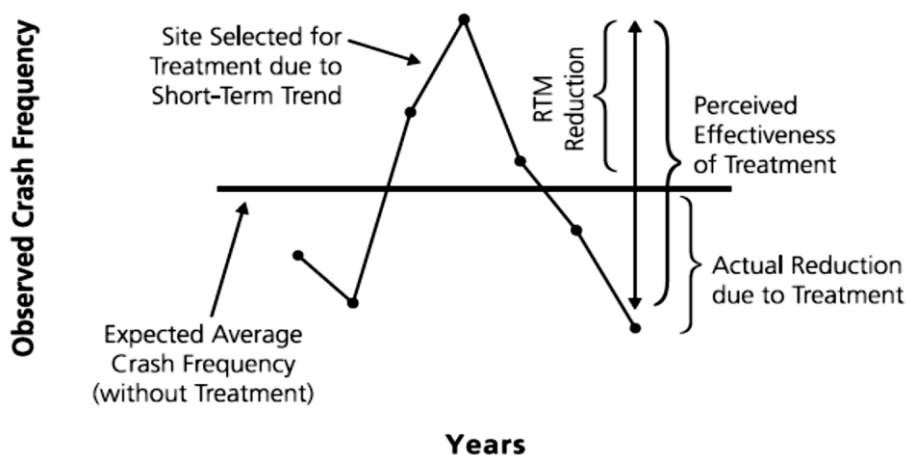


Figura 2.1: Regressão à média no “antes” e “depois”. Fonte: Manual de Segurança Rodoviária (HSM)

Alguns trabalhos de investigação neste campo têm mostrado que este fenómeno pode produzir reduções aparentes na frequência dos acidentes na ordem dos 5 a 30% (Wright e Boyle, 1987), sendo que os sítios sujeitos a elevadas frequências num curto espaço de tempo tendem a ser mais sensíveis ao seu efeito. Nesse sentido, é importante garantir a aplicação de métodos que tenham em consideração o fenómeno de regressão à média. É o caso da aplicação de modelos de previsão de acidentes, por exemplo, na deteção de ZAAs.

2.3. MODELOS DE FREQUÊNCIA DE ACIDENTES

Os acidentes têm características específicas, que se ajustam a uma análise com base na aplicação de modelos econométricos, dado que os acidentes são acontecimentos aleatórios. O facto de os acidentes serem eventos aleatórios e imprevisíveis, não é limitativa para que o número de acidentes seja objeto de uma análise causal ou de intervenção política (OCDE, 1997).

Os MPAs são modelos estatísticos que permitem determinar uma frequência esperada de acidentes num determinado local, assim como, identificar fatores que contribuem para a ocorrência dos mesmos, através de variáveis que refletem as condições geométricas e funcionais do local considerado (geralmente, segmentos ou interseções) (Claude, 2012).

Uma vez que a frequência de acidentes ou de vítimas corresponde a um número inteiro positivo, o modelo matemático a selecionar deverá ser escolhido, tendo por base esta característica para a variável dependente (ou variável de resposta do modelo).

Os modelos que apresentam uma aplicação mais alargada para a previsão de acidentes, segundo Lord e Mannering (2010) são: Regressão de Poisson, Regressão Binomial Negativa, Regressão Poisson Lognormal, Regressão Inflacionada de Zero, Modelo de Parâmetros Aleatórios, Modelo Multinomial Negativo, Regressão Gama, Equações de Estimação Generalizadas, Modelo Multivariado, Modelos hierárquicos e Modelos de *Machine Learning*.

Contudo, os mais comumente aplicados e que serão utilizados no presente estudo são: o modelo de Regressão de Poisson e Regressão Binomial Negativa.

2.3.1. MODELOS DE REGRESSÃO DE POISSON E BINOMIAL NEGATIVO

A distribuição de Poisson é uma distribuição probabilística, que descreve a probabilidade do número de ocorrências (variável aleatória) num intervalo contínuo (tempo ou espaço). Deste modo, esta unidade de medida é um número inteiro e positivo.

A regressão de Poisson baseia-se na análise de frequências e apresenta uma propriedade em que a variância é igual à média, ou seja, uma vez estimado o valor esperado, sabe-se o que esperar relativamente à variação da média. Esta distribuição especifica que a observação y_i tem uma probabilidade de ocorrer segundo a distribuição probabilística de Poisson. É caracterizada pelo parâmetro λ_i que, por sua vez, é função de um conjunto de variáveis independentes x_i (Greene, W. H., 2008). A formulação que caracteriza a distribuição de Poisson é:

$$P[Y = y_i] = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \quad (2.1)$$

Assim, é caracterizada por uma média $\mu_i = E[Y] = \lambda_i$ e variância $Var[y_i] = \lambda_i$, ou seja, a média é igual à variância.

A relação entre o parâmetro de Poisson λ_i e as variáveis independentes é estabelecida através da função logarítmica de ligação:

$$\ln \lambda_i = \sum_j \beta_j x_{ij} \quad (2.2)$$

Esta função apresenta β_j como o parâmetro a ser estimado para as variáveis independentes x_{ij} , sendo j o número de variáveis independentes.

Contudo, a homogeneidade do risco individual de cada entidade, de um conjunto homogêneo de entidades, é posta em causa por vários fatores, resultando valores de variância diferentes do valor esperado que são iguais ao valor médio (Cardoso, J., 2007).

No caso da frequência de acidentes, é comum que a variância suplante a média para os dados de acidentes, apresentando o modelo uma sobredispersão. Deste modo, a propriedade da distribuição de Poisson (média igual à variância) não é obedecida e pode resultar em estimativas de parâmetros enviesados que, por sua vez, levam a conclusões incorretas sobre os fatores que determinam as frequências de acidentes (Maycock and Hall, 1984; Maher and Summersgill, 1996)

De forma a superar este problema, recorre-se à distribuição binomial negativa (BN), que tem vindo a ser mais frequentemente aplicada, uma vez que considera que os acidentes resultam de

um processo cujas probabilidades de sucesso variam em cada evento. Esta distribuição trata-se de uma extensão da distribuição de Poisson, sendo mais adequada para a modelação dos dados de acidentes.

Na distribuição BN, caso se verifique sobredispersão, as estimativas serão quase-verosimilhanes (Cardoso, J., 2007). A forma funcional que caracteriza a distribuição BN é representada pelas seguintes equações:

$$\Pr[Y = y_i] = \frac{\Gamma(\theta + y_i)}{\Gamma(\theta)\Gamma(y_i + 1)} r_i^\theta (1 - r_i)^{y_i} \quad (2.3)$$

$$r_i = \lambda_i / (\theta + \lambda_i) \quad (2.4)$$

$$\theta = 1/\alpha \quad (2.5)$$

$$\text{Var}[y_i] = E[y_i]\{1 + \alpha E[y_i]\} \quad (2.6)$$

Assim, o modelo de regressão binomial negativo é adequado para modelar dados sobredispersos, uma vez que incorpora um parâmetro estatístico adicional de sobredispersão, que é estimado em conjunto com os parâmetros da equação de regressão. Para além deste parâmetro ser positivo, quanto maior for, maior é a heterogeneidade dos locais em estudo.

O parâmetro de sobredispersão permite avaliar a capacidade de o modelo explicar a componente sistemática do processo de ocorrência de acidentes. Dado que uma das razões deste fenómeno é a omissão de variáveis na componente sistémica, quanto menor o valor do parâmetro, melhor é o modelo.

Considera-se que o valor esperado da frequência de acidentes, a longo prazo, resulta de um processo causal. A atribuição da variação de fatores causais é sistémica e pode, à partida, ser controlada e influenciada. Assim, esta componente apresenta um elevado interesse de estudo, para efeitos de estratégia política da segurança rodoviária (OCDE, 1997).

Nesse sentido, uma componente importante no desenvolvimento de modelos de previsão da frequência de acidentes é a determinação da relação entre as variáveis independentes e o número de acidentes ou de vítimas (variáveis dependentes), através de funções matemáticas. Deste modo,

é possível quantificar a variação da sinistralidade relacionada com a alteração de cada variável independente considerada, designando-se de elasticidade (Gomes, 2010).

Maher, M. and Summersgill, I. (1995) referem que os princípios de escolha das variáveis devem ter em conta quatro fatores:

- Nível de significância – as variáveis que apresentem níveis de confiança inferiores a 95% devem ser rejeitadas e, para níveis de confiança superiores a 99%, apenas serão rejeitadas, caso o estudo assim o justifique;
- Estabilidade do modelo – dada a relação entre as variáveis, caso ocorra a introdução de uma nova variável, os valores dos parâmetros das outras variáveis presentes no modelo, serão afetados;
- Efeito da variável deve ser lógico – a estrutura do modelo e o valor do parâmetro ajustado aos conhecimentos empíricos devem fazer sentido;
- Facilidade de medição da variável e dimensão do efeito – as variáveis que tenham um efeito preponderante e que sejam mais fáceis de medir devem ser preferenciais.

Sintetizando, segundo McCullagh, P. and Nelder, J. A. (1989), os modelos podem ser especificados em três partes:

- a variável dependente (ou resposta): componente aleatória do modelo (Y) - do qual resulta um conjunto de n observações y - independente da distribuição Normal em que $E[Y] = \mu$ e a variância constante σ^2 ;
- as variáveis independentes: entram na forma de uma estrutura linear, constituindo o componente sistemático do modelo ($x_1, x_2, \dots, x_p$);
- a ligação entre os componentes aleatórios e sistemáticos é feita através de uma função adequada como, por exemplo, logarítmica para os modelos log-lineares, chamada função de ligação.

Variáveis Independentes

A escolha das variáveis independentes é uma etapa importante para a modelação e depende do objetivo da aplicação do modelo. É necessária a existência de correlação significativa entre as variáveis independentes, e a variável dependente de interesse. Assim, a criação da base de dados é um fator importante para a determinação do MPA.

Como referido por Madalozo (2003), o número de acidentes pode estar diretamente relacionado com as variáveis do traçado em planta e em perfil longitudinal, assim como da secção transversal da rede rodoviária. As variáveis independentes usadas, em geral, nos MPA, e que permitem descrever fatores de risco, são de dois tipos: variáveis de exposição ao risco, variáveis que caracterizam a geometria e funcionalidade do local. (Fridstrom, L., et al., 1995; Lord, D., 2000; Greibe, P., 2003; Golob, T., et al., 2004; Polus, A., et al., 2005; Reurings, M., et al., 2005).

É de salientar que, na modelação de frequência de acidentes por local (segmento ou intersecção), os dados são agrupados, o que implica que os dados individuais dos acidentes não sejam incluídos. Deste modo, os fatores relacionados com o veículo, condutor, condições meteorológicas, etc, não são expressos neste tipo de modelação. Como Kockelman e Kweon (2002) referem, uma vez que os dados são associados pelo local e as características individuais dos acidentes não podem ser integradas, estes fatores perdem relevância no modelo.

Relativamente às principais variáveis independentes utilizadas nos modelos de frequência de acidentes e de gravidade, estas estão relacionadas com fatores viário-ambientais (Das & Abdel-Aty, 2011; Zeng et al., 2016b). Dentre os modelos de frequência de acidentes com gravidade, as variáveis mais importantes a serem consideradas são: velocidade máxima regulamentar, volume de tráfego, tipo de separador, pavimento, largura da via e número de vias. É ainda de notar, a importância do volume de tráfego, extensão do segmento, alinhamento horizontal e velocidade máxima regulamentar nos modelos de previsão de frequência de acidentes por segmento.

As variáveis independentes devem ser selecionadas com base em estudos prévios, para que sejam medidas e determinadas de forma fiável e não apresentem correlação com outras variáveis previamente incluídas (Reurings, M., et al., 2005). Conforme Fridstrom, L., et al. (1995), o método passa pela introdução das variáveis independentes uma a uma, incluindo apenas no modelo as que apresentam significância estatística com a variável dependente em estudo. Na prática, a escolha das variáveis é, maioritariamente, limitada pelos dados disponíveis.

Função do modelo

Como referido por Lord e Mannering (2010), a forma funcional do modelo estabelece a relação entre a variável dependente e a variável independente, tornando-se uma parte crítica no processo de modelação, uma vez que o efeito não linear que as variáveis independentes apresentam na frequência de acidentes, não pode ser representado. Deste modo, a maioria dos modelos de contagem de dados de acidentes assumem que as variáveis independentes apresentam, de alguma forma, uma influência linear nas variáveis dependentes. Contudo, existem estudos que sugerem que a função não linear caracteriza melhor a relação entre a frequência de acidente e as variáveis independentes. No entanto, a função não linear pode apresentar alguma complexidade e exigir procedimentos de estimativas (Miaou and Lord, 2003; Bonneson and Pratt, 2008).

Assim, a formulação mais conhecida e utilizada atualmente, no que se refere à análise multivariada é o modelo de regressão linear:

$$\eta_i = \sum_i \beta_i x_{ji} \quad (2.7)$$

Na função, i corresponde ao número de observações, j ao número de variáveis independentes, η_i é a componente sistemática do modelo, β_j os parâmetros a estimar e x_{ji} as variáveis independentes. É de notar que as variáveis x_{ji} podem assumir qualquer forma, isto é, o modelo é linear para os parâmetros β_j , não sendo necessariamente nas variáveis x_{ji} .

Dado que o valor esperado de acidentes tem de ser positivo, é importante garantir esta condição na função, considerando por exemplo:

$$E[Y] = e^{\sum_j \beta_j x_j} \quad (2.8)$$

Esta função encontra-se linearizada através da função logarítmica.

Atualmente, a forma mais utilizada é (Reurings, M., et al., 2005; Cardoso, J., 2007):

$$E[Y] = \alpha Q^\beta e^{\sum_j \gamma_j x_j} \quad (2.9)$$

$E[Y]$ corresponde ao valor esperado da frequência de acidentes e é uma função multiplicativa do tráfego – Q , e de um conjunto de p fatores de risco $x_j = x_1, x_2, x_3, \dots, x_p$, em que os parâmetros a estimar são α , β e γ_j .

A variável volume de tráfego (normalmente o TMDA) apresenta uma relação com a frequência de acidentes, dada pela função potência, sendo a função exponencial a forma mais utilizada para relacionar a frequência dos acidentes com as variáveis x_j . A função potência apresenta duas características fundamentais. Por um lado, representa a situação em que quando não há tráfego, não ocorrem acidentes. Em contrapartida, estabelece uma relação não linear entre os acidentes e o fluxo de tráfego (Lord, D., 2000).

No caso de o valor β tomar valor de 1, o valor esperado da frequência dos acidentes é proporcional ao volume de tráfego. Contudo, é comum que na maioria dos modelos o valor de β seja diferente da unidade, possivelmente, devido ao facto de não se desagregar a variável dependente por tipos ou a alterações das características da circulação do tráfego (Cardoso, J., 2007).

Estimativa dos parâmetros da função

Como referido anteriormente, a dispersão advém, principalmente, devido à natureza aleatória do processo de ocorrência de acidentes, uma vez que estes resultam de um processo de ensaios de Bernoulli com probabilidades distintas, ou seja, locais iguais com as mesmas características e com valores da média diferentes entre si (Lord, D., 2006; Lord, D. and Park, P. Y.J., 2008).

Quando se trata do tipo de regressão de Poisson e BN (regressões não lineares), a estimativa dos parâmetros comumente aplicada é o método de máxima verosimilhança. Este método consiste em seleccionar, entre todos os valores possíveis dos parâmetros, aqueles que representam de forma mais verosímil a ocorrência de uma amostra semelhante àquela que se obteve (Guimarães, R. C. and Cabral, J. S., 1997). Deste modo, o autor menciona que a probabilidade de ocorrência dos valores observados de uma amostra aleatória simples idêntica à que se obteve é dada pela equação 2.10:

$$L = Probabilidade(Y_1 = y_1, \dots, Y_N = y_N | \beta_1, \beta_2, \dots, \beta_R) \quad (2.10)$$

As estimativas de máxima verosimilhança dos parâmetros $\beta_1, \beta_2, \dots, \beta_R$ são os valores destes parâmetros, que maximizam a função de verosimilhança. Segundo Guimarães, R. C. e Cabral, J. S. (1997), estas podem ser obtidas resolvendo o sistema de equações:

$$\frac{\partial L}{\partial \beta_i} = 0 \quad (i = 1, \dots, R) \quad (2.11)$$

Caso se trate da regressão de Poisson, a função logarítmica de verosimilhança é (Greene, W. H., 2008):

$$\ln L = \sum_{i=1}^n [-\lambda_i + y_i x_i \beta - \ln y_i!] \quad (2.12)$$

O processo de estimativa dos parâmetros associado a estas distribuições implica a resolução de sistemas de equações não lineares, iterativamente até que os critérios de convergência sejam satisfatórios (Cardoso, J., 2007).

Com o objetivo de estudar melhor a influência dos resultados da estimação dos parâmetros, calcula-se a elasticidade para determinar o efeito que as variáveis independentes apresentam na variável resposta. Esta corresponde a uma estimativa do impacto de uma variável no valor percentual esperado da frequência dos acidentes, ou seja, o efeito no valor esperado da frequência dos acidentes μ_i resultante de uma alteração de 1% na variável.

Assim, a elasticidade de uma variável contínua é explicada através da equação (Washington, S., et al., 2003):

$$E_{x_{ij}}^{\mu_1} = \frac{\partial \mu_i}{\mu_i} x \frac{x_{ij}}{\partial x_{ij}} = \beta_j x_{ij} \quad (2.13)$$

Na qual E corresponde à elasticidade, x_{ij} é o valor da variável independente j para a observação i , β_j é o parâmetro a estimar da variável j e μ_i é o valor esperado da frequência de acidentes para a observação i .

Caso as variáveis independentes x_j sejam “dummys”, isto é, não sejam variáveis contínuas e tomem valores 0 ou 1, os seus parâmetros medem aproximadamente o efeito relativo no valor esperado $E[Y]$ e têm como designação pseudo-elasticidade (Fridstrom, L., et al., 1995; Washington, S., et al., 2003):

$$E_{x_{ij}}^{\mu_1} = \frac{e^{\beta_j} - 1}{e^{\beta_j}} \quad (2.14)$$

Em suma, para se desenvolver um MPA é necessário:

1. Definir as variáveis independentes x ;
2. Determinar a função f ;
3. Definir a distribuição probabilística que rege a componente aleatória;
4. Estimar os parâmetros da função $f(x)$.

2.3.2. SEGMENTAÇÃO DA REDE

Esta etapa constitui o desafio primário na previsão de frequência de acidentes, uma vez que a escolha da estratégia a adotar para a segmentação da rede viária tem impacto nos resultados obtidos e respetiva interpretação. No caso da rede de autoestradas, como não existem interseções de nível, a segmentação é definida pela extensão do segmento a considerar. A segmentação pode ser de dois tipos: valor da extensão do segmento homogéneo, isto é, não varia de segmento para segmento; ou valores de extensão variáveis, e definidos de acordo com as características geométricas do segmento. O primeiro caso utiliza-se quando há uma variação das características geométricas (planta, perfil longitudinal e transversal) muito elevada, o que dificulta o processo de segmentação. Neste caso, é necessário calcular índices para algumas das variáveis, de forma a uniformizar a variabilidade da variável, no mesmo segmento (por exemplo, o segmento pode ter duas vias numa parte da sua extensão, e três vias noutra extensão do mesmo segmento). Este procedimento dificulta a interpretação do efeito de algumas variáveis. No caso das autoestradas, em geral, as características geométricas são mais uniformes ao longo de um determinado sublanço, o que possibilita definir segmentos com extensão variável de acordo com a transição das suas características ao longo do sublanço. Contudo, tanto segmentos muito extensos, como muito curtos, são inadequados.

Por um lado, um segmento muito longo implica a escassez de informação detalhada. Por outro lado, quando um segmento é pequeno, implica mais variações, que introduzem incertezas no modelo (Cook *et al.*, 2011; Lu *et al.*, 2013; Green *et al.*, 2017). Quantas mais variáveis forem utilizadas para avaliar a homogeneização do segmento, mais curtos estes serão (Resende & Benekohal, 1997). Estes segmentos curtos implicam uma redução da robustez estatística dos MPA. Além disso, a um grande número de segmentos, estão associados mais segmentos com zero acidentes, devido a características de aleatoriedade inerente aos acidentes (Souleyrette *et al.*, 2007).

Deste modo, não existe consenso no que diz respeito aos valores mínimos e máximos da extensão do segmento, dado que o processo de segmentação é, maioritariamente, dependente dos atributos escolhidos e do foco da análise de segurança (Cafiso *et al.*, 2008; Borsos *et al.*, 2014). Ainda, é necessário ter em conta a precisão da localização dos acidentes, uma vez que esta influencia a escolha do tamanho do segmento (Green e Agent, 2011; Qin e Wellner, 2012).

A abordagem adotada neste trabalho resulta de um conjunto de elementos viários individuais, de comprimento variável, apresentando homogeneidade em algumas características, tais como: geometria e volume de tráfego.

Segundo o *Highway Safety Manual* (HSM), para a realização dos MPAs, deve-se ter em atenção alguns critérios para avaliar a homogeneidade e divisão dos segmentos. Nas autoestradas é importante ter por base atributos como: largura e tipo de separador, largura da berma, largura da via, número de vias, vias auxiliares, raio de curvatura, comprimento da reta, volume de tráfego, limite de velocidade e controlo de velocidade.

2.4. VALIDAÇÃO

A validação constitui um papel importante no desenvolvimento de modelos e tem como objetivo averiguar a adequabilidade e utilidade do modelo ao propósito para o qual foi desenvolvido, aumentando a confiança quanto aos resultados obtidos (Power, 1993; Oh *et al.*, 2003; Biondi *et al.*, 2012). Esta etapa permite avaliar de uma forma precisa o modelo, assim como, identificar mais corretamente as suas limitações, possibilitando a sua correção para o desenvolvimento de modelos melhores (Biondi *et al.*, 2012).

Nesta etapa é possível verificar se o modelo capturou adequadamente a relação subjacente nos dados utilizados para calibração do modelo e permite através da análise de variáveis e parâmetros, que avaliam o comportamento do modelo em relação ao sistema real, atribuir fiabilidade ao mesmo.

Deste modo, são apresentadas as medidas de desempenho utilizadas para verificação da validade dos modelos quanto à capacidade de previsão, assim como as análises gráficas selecionadas para complementar, tendo por base o estudo de validação aplicado por Silva, P (2019). Para as medidas de desempenho temos: y_i – é o valor observado; $\hat{y}_i$ – é o valor previsto; $\bar{y}_i$ – é a média dos valores observados.

2.4.1. MEDIDAS DE DESEMPENHO

Erro Relativo

Esta métrica relaciona o desvio não explicado e o desvio total, sendo que o desvio corresponde à diferença entre o valor estimado e o valor observado num determinado local. É um índice relativo em que, quanto menor o erro, menor é a parcela de variação não explicada em relação à variação total. Assim, quanto mais próximo de zero, melhor o desempenho do modelo. Acresce a isso que:

- ER=1: significa que o modelo desempenha tão bem quanto a simples média dos valores observados;
- ER > 1: significa que a média observada é melhor preditor que o modelo.

$$ER = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.15)$$

Coeficiente de determinação (R^2)

O coeficiente de determinação, R^2 mede o ajuste de um modelo estatístico linear generalizado, relativamente aos valores observados de uma variável aleatória. É uma boa métrica para avaliar o ajuste dos dados ao modelo de regressão. O coeficiente de determinação estabelece a relação entre a soma dos quadrados da diferença entre a média das observações e o valor estimado para cada observação e a soma dos quadrados das diferenças entre a média e cada valor observado. Quanto mais próximo de 1, melhor o modelo se ajusta à amostra. Assim:

- correlação perfeita diz respeito a $R^2=1$ – indica que os dados observados são idênticos aos previstos.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.16)$$

Estatísticas Resumidas

A análise dos conjuntos de dados observados e previstos, através da comparação entre as estatística resumidas (mínimo, máximo, média e desvio padrão), permite a verificar a aproximação dos resultados, dada a comparação das características como um todo e não individualmente (Benett et al., 2013). Nesse sentido, quando os valores das estatísticas calculadas para as previsões se aproximam dos valores das estatísticas observadas, o modelo apresenta um bom ajuste (Silva, P. 2019).

$$\begin{aligned}
 Min_{obs} &= \min(y_1, y_2, \dots, y_n) \\
 Máx_{obs} &= \max(y_1, y_2, \dots, y_n) \\
 \bar{y}_i &= \frac{\sum_{i=1}^n y_i}{n} \\
 DP &= \frac{\sqrt{\sum_{i=1}^n (y_i - \bar{y}_i)^2}}{n-1}
 \end{aligned}
 \tag{2.17}$$

Erro Logarítmico Quadrático Médio (MSLE)

Este índice é uma variação do erro médio absoluto e determina o logaritmo dos valores reais e previstos, relacionando a diferença entre os mesmos. Dada a sua natureza logarítmica de erro, o MSLE tem atenção às diferenças percentuais. Isso significa que trata pequenas diferenças entre pequenos valores da mesma forma que grandes diferenças em grande escala. Idealmente este valor deve ser próximo de zero, o que significa que o valor previsto é próximo do real.

$$MSLE = \frac{1}{n} \sum_{i=1}^n (\ln y_i - \ln \hat{y}_i)^2
 \tag{2.18}$$

Para aplicar o logaritmo, é necessário garantir que os valores observados e previstos sejam diferentes de zero. Assim, adiciona-se o valor de 1 a todas as observações, tanto para o valor observado, como para o valor previsto. Deste modo, a função pode ser utilizada, se o valor observado ou o previsto tiverem elementos iguais a zero, como acontece na maior parte dos casos (Ben Hamner et al., 2018).

Desvio Absoluto Médio (MAD)

O desvio absoluto médio de um conjunto de dados é a distância média entre cada valor e a média da amostra, descrevendo as variações num conjunto de dados. O MAD é uma alternativa proposta para utilização, como substituto do desvio padrão, uma vez que é mais fiel à realidade. Ainda, permite perceber como é que os valores de um conjunto de dados foram "divididos". A precisão da previsão deste método está intimamente relacionada com o método do erro quadrático médio (MSE). Embora estes métodos estejam relacionados, o MAD é o mais utilizado, uma vez que é mais fácil de calcular (evitando a necessidade do quadrado), logo mais simples de compreender.

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

(2.19)

Erro Percentual Absoluto Médio Simétrico (SMAPE)

O SMAPE é uma medida de ajuste com base em erros percentuais (ou relativos). A referência mais antiga à fórmula diz respeito a Armstrong (1985, p. 348), e é representada como "MAPE ajustada", definida sem os valores absolutos no denominador. Posteriormente, esta foi discutida, modificada e reproposta por Flores (1986).

O erro percentual absoluto médio simétrico é definido como duas vezes o somatório da diferença absoluta entre o valor observado e previsto, dividido pelo somatório da adição dos valores absolutos do valor observado e o valor previsto. Este cálculo é posteriormente dividido pelo total de observações. Assim, apenas será impossível calcular esta função caso, tanto o valor observado, como o previsto, forem zero. Esta métrica tem como limite superior o valor 2, quando o valor observado ou o previsto é zero, ou quando o valor observado e o previsto tem sinais opostos.

$$SMAPE = \frac{2}{n} \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\sum_{i=1}^n |y_i| + |\hat{y}_i|}$$

(2.20)

Erro Quadrático Médio de Previsão (MSPE)

O MSPE fornece uma descrição simples sobre o desempenho do modelo. Este índice corresponde ao valor da diferença ao quadrado entre os valores observados e os previstos, a dividir pelo total de observações. Idealmente, este valor deveria ser próximo de zero, o que significa que o seu preditor está próximo do valor real.

$$MSPE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (2.21)$$

2.4.2. ANÁLISE GRÁFICA

Análise de Resíduos Acumulados

O CURE (*CUmulative REsiduals*) é um método proposto por Hauer e Banfo (1997), que se baseia na análise de resíduos (diferença entre valores observados e os previstos) acumulados, permitindo analisar a variação da linha e distinguir o que é um padrão que representa um bom ajuste do modelo e o que é um padrão de enviesamento sistemático do modelo.

Este processo requer a pré-estimação de parâmetros, realizada na etapa da modelação, para determinar valores previstos e, com estes, os resíduos de cada observação. De seguida, o gráfico é realizado com os acumulados desses resíduos, ordenados por ordem crescente, de acordo com a variável independente determinada para a análise. É desejável que os resíduos acumulados tenham um comportamento gráfico, no qual, oscilem em torno de zero e não extrapolem as curvas de limites, consideradas aceitáveis, para os resíduos acumulados. Considere-se que N é o número total de resíduos, n é um inteiro entre 1 e N , obtém-se as curvas através de:

- Disposição dos resíduos acumulados por ordem crescente da variável de interesse, numerando consecutivamente 1,2, ..., j,...,N;
- Cálculo do quadrado dos resíduos ordenados;
- Cálculo $\sigma^2(j)$, como soma dos resíduos quadrados de 1 a j;
- Cálculo $\sigma'(j) = \sigma(j) \sqrt{1 - \frac{\sigma^2(j)}{\sigma(N)}}$;
- Determinação dos limites $\pm 2\sigma'(j)$.

A partir deste método é possível avaliar a qualidade de ajuste do modelo de previsão determinado, e a influência das variáveis independentes, que vão sendo inseridas no mesmo. Realizando uma análise gráfica, podemos observar se o modelo é satisfatório, ou seja, se os resíduos acumulados, ordenados de acordo com o valor da variável independente, oscilam em torno da origem, e terminam com um valor final acumulado próximo de zero, não ultrapassando o limite determinado pelos valores $\pm 2\sigma'$.

Gráfico de Dispersão de Valores Observados e Previstos

Este método compara diretamente os valores observados e previstos. Trata-se de uma estratégia prática para identificar, rapidamente, como as previsões podem diferir das observações, o que não é muito óbvio, analisando as estatísticas resumidas. Idealmente, o comportamento deveria ser de concordância pontual entre os dois dados distintos e de tendência geral dos dados.

Em resumo, o procedimento seguido nesta etapa, para os modelos desenvolvidos, foi:

- Cálculo das medidas de desempenho (métricas de erro e R^2);
- Construção de gráficos dos resultados de previsão e resíduos;
- Comparação entre as medidas de desempenho como passo inicial da seleção dos melhores modelos;
- Análise dos gráficos de resultados para confirmar a validade dos modelos.

3

CASO DE ESTUDO

3.1. METODOLOGIA

A modelação da segurança rodoviária pode fornecer dois tipos de informação, segundo Kim e Washington (2006) e Hauer (2004): a estimativa da frequência esperada de acidentes de um elemento da infraestrutura, com base nas suas características e a estimativa da influência das características de um elemento da infraestrutura, na frequência esperada de acidentes. Este trabalho tem como objetivo a primeira referência.

Para formular um modelo linear generalizado é necessário escolher:

- i. uma distribuição de probabilidade para a variável dependente, que deve fazer parte das distribuições exponencial;
- ii. as variáveis explicativas, que podem ser quantitativas e/ou qualitativas;
- iii. uma função de ligação que irá relacionar a componente aleatória e a sistemática do modelo.

Dado que a modelação tem como objetivo o ajuste de um modelo aos dados de acidentes, características geométricas e operacionais de uma via (ou segmento), representando a realidade de uma forma simples, é importante ter em conta os fatores mais preponderantes para a sua descrição (Hauer, 2004; Chang, 2005). Deste modo, a natureza dos dados a serem tratados é o primeiro passo e tem uma elevada importância.

Neste capítulo foi desenvolvido o modelo de previsão da frequência de acidentes para a autoestrada A44 da rede ASCENDI, focando a análise na estrada em plena via, denominada por sublanço. Este estudo foi realizado através da aplicação dos modelos Poisson e BN, desenvolvendo um modelo de previsão quantitativo, que assenta no conceito da utilização histórica dos dados dos acidentes, para estimar os acidentes futuros.

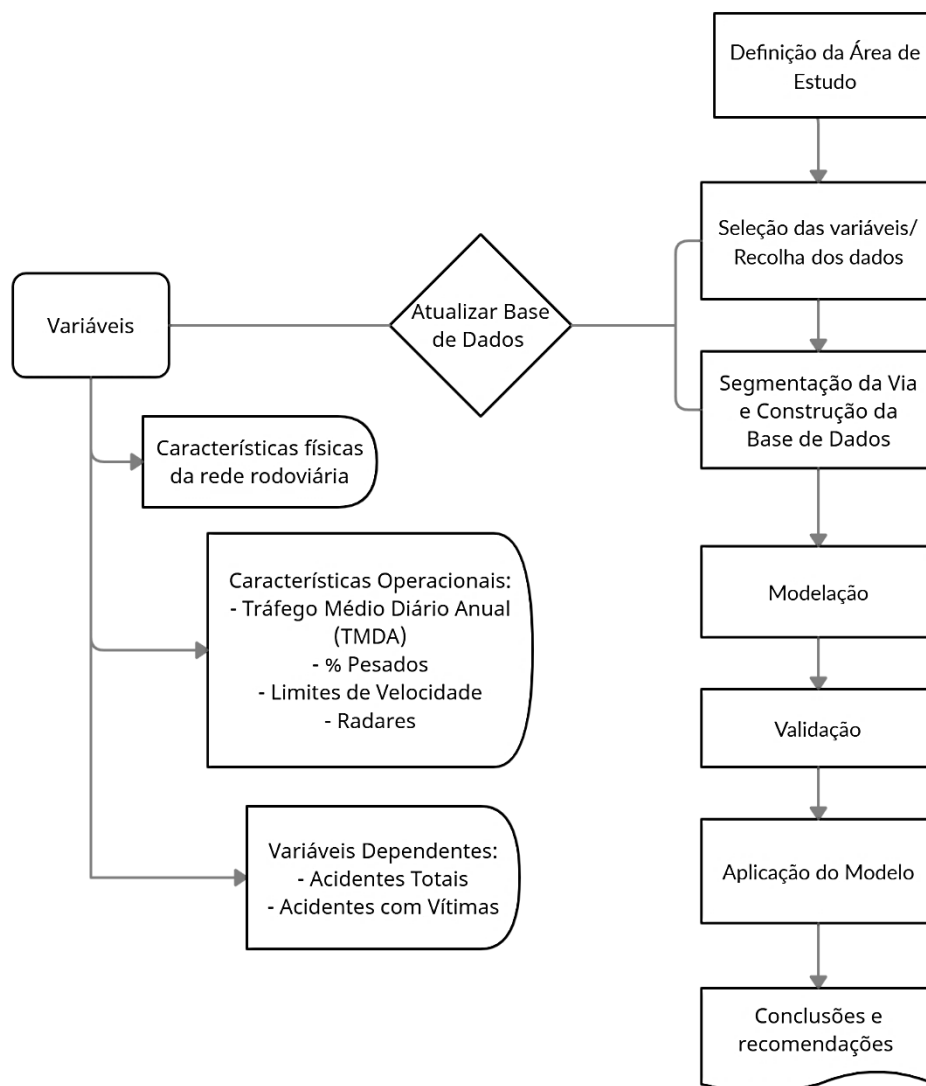


Figura 3.1: Estrutura da metodologia

3.2. ENQUADRAMENTO DO CASO DE ESTUDO

A primeira etapa do processo passa por definir a área de estudo pretendida, tendo em conta a disponibilidade de dados. É necessário considerar critérios para a escolha da autoestrada, como as características técnicas e operacionais dos trechos da rede viária, assim como a sinistralidade.

A Ascendi detém participações maioritárias em seis concessões portuguesas e um contrato de prestação de serviços de Operação e Manutenção na subconcessão do Douro Interior, apresentando um rede com extensão de 868km - figura 3.2.

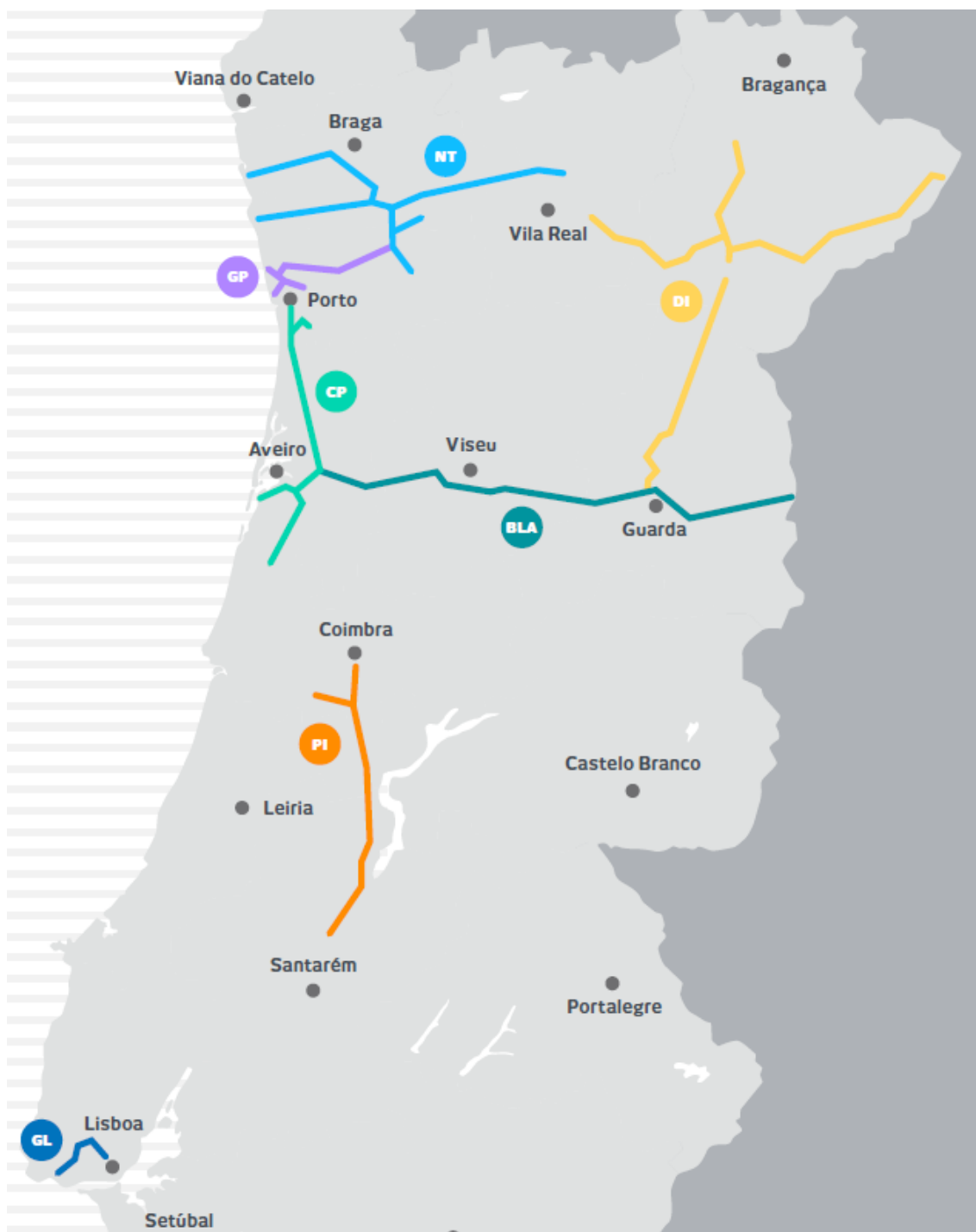


Figura 3.2: Concessões Ascendi. Fonte: Anuário de Segurança Rodoviária Ascendi, 2020

Podemos observar no quadro 3.1 os dados referentes à rede rodoviária da Ascendi:

Quadro 3.1: Concessões Ascendi

Concessões/ Subconcessões	Autoestradas	Extensão (Km)
Norte (CN)	A7, A11	179
Costa de Prata (CP)	A17, A25, A29, A44	105
Grande Porto (GP)	A4, A41, A42, VRI	55
Beira Litoral e Alta (BLA)	A25	172
Grande Lisboa (GL)	A16	23
Subconcessão Pinhal Interior (SPI)	A13, A13-1	93
Subconcessão Douro Interior (SDI)	IP2, IC5,	242

Os dados fornecidos pela Ascendi foram analisados de forma a identificar e extrair informação necessária para caracterizar as variáveis a utilizar, com vista ao desenvolvimento dos MPA.

Numa primeira instância e após a identificação da rede rodoviária, foi realizado uma breve análise de sinistralidade de toda a rede Ascendi, com o objetivo de identificar as autoestradas que apresentam maior importância para o estudo de segurança rodoviária. A análise da base de dados tem como foco os dados fornecidos de acidentes totais e acidentes com vítimas, relativos aos anos 2015, 2016, 2017, 2018, 2019 e 2020.

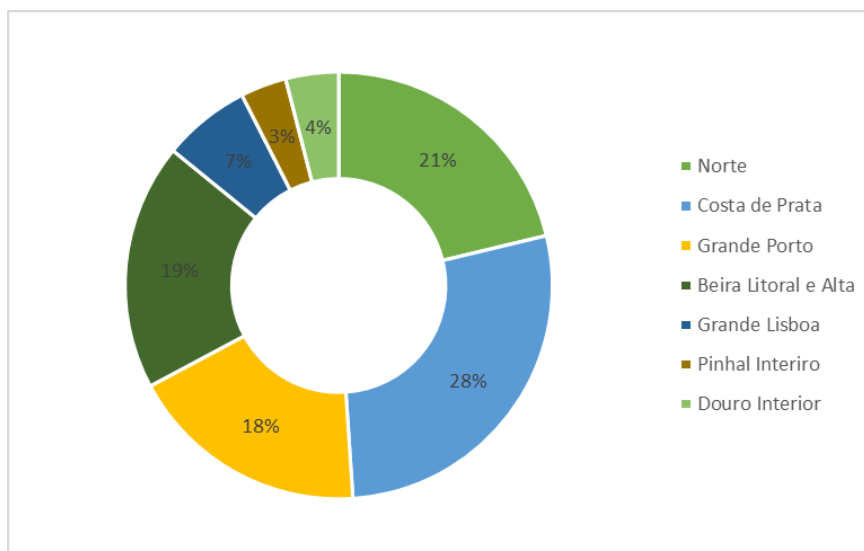


Figura 3.3: Percentagens de Acidentes Totais por Concessão [2015-2020]

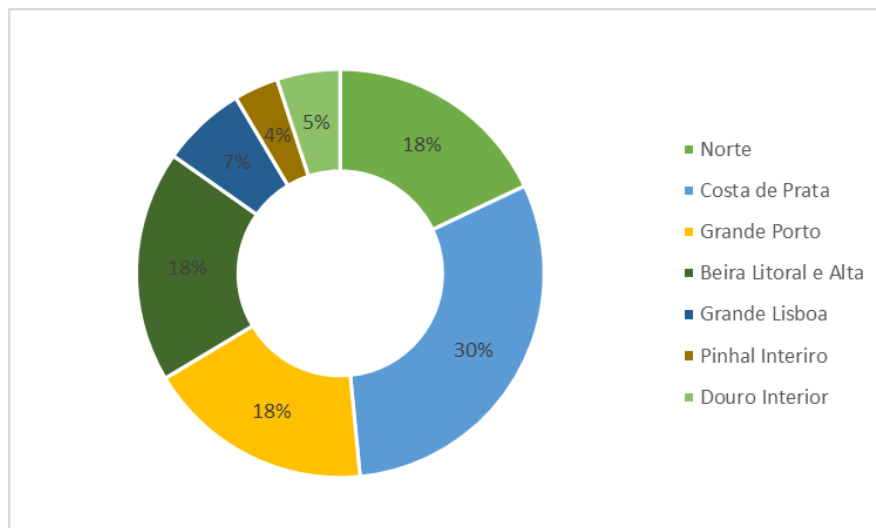


Figura 3.4: Percentagens de Acidentes com Vítimas por Concessão [2015-2020]

Através da análise gráfica, comprova-se que a concessão com maior taxa de sinistralidade é a da Costa de Prata (28%), seguindo-se a concessão Norte (21%). No que diz respeito a acidentes com vítimas, podemos constatar que a Costa da Prata se mantém em primeiro lugar (30%), seguindo-se a concessão Norte e Grande Porto, que apresentam o mesmo valor nesta classificação (18%).

Deste modo, é realizada a exploração das autoestradas pertencentes à concessão da CP, com o objetivo de estabelecer uma hierarquia de importância. No gráfico da figura 3.5, pode observar-se os valores relativos aos dados de acidentes totais e acidentes com vítimas, assim como a relação da sinistralidade com a extensão, das autoestradas A17, A25, A29 e A44:

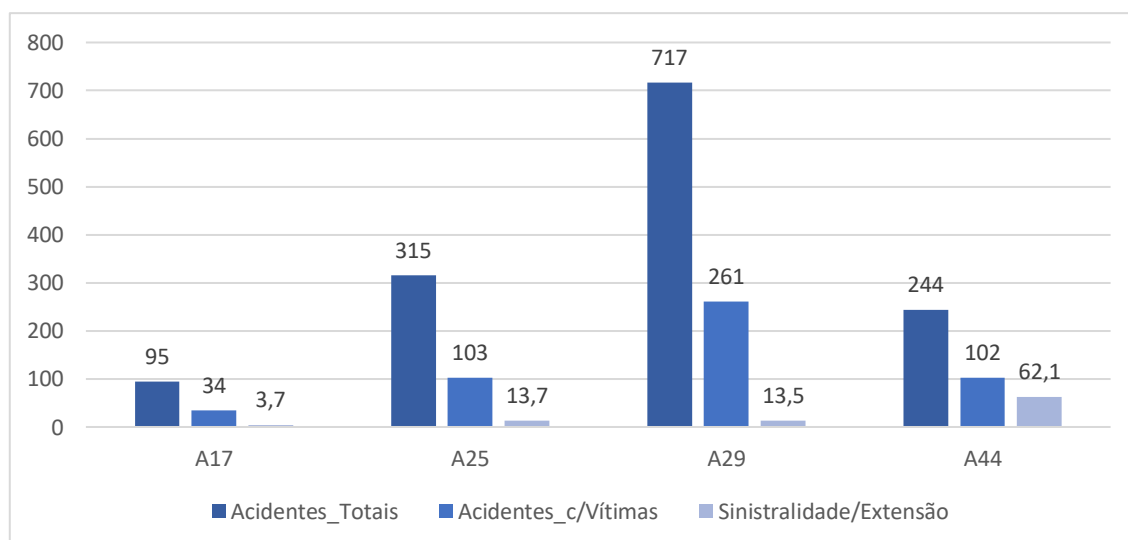


Figura 3.5: Gráfico de acidentes totais, acidentes com vítimas e relação sinistralidade/extensão, nas autoestradas da CP [2015-2020]

A partir do gráfico é possível observar que a autoestrada A29 destaca-se com uma elevada taxa de sinistralidade, relativamente às outras. A autoestrada A25 apresenta-se em segundo lugar na classificação de acidentes totais, com 315, seguindo-se a A44 com 244.

Considerando o trabalho realizado anteriormente por Souza, S. (2021), e que teve como objetivo o desenvolvimento de um MPA para a autoestrada A29, pretende-se dar continuidade ao estudo anterior, aplicando neste caso à autoestrada A44. Esta autoestrada foi selecionada, devido à sua elevada sinistralidade para uma pequena extensão, relativamente às autoestradas pertencentes à concessão Costa de Prata. De facto, apesar da autoestrada A25 apresentar valores superiores para o número total de acidentes, a relação entre frequência de acidentes por km é menor do que na A44, ou seja, 13,8 e 62,1, respetivamente.

3.2.1. DEFINIÇÃO DA ÁREA DE ESTUDO

Com o objetivo de estudar as diferentes autoestradas da rede Ascendi de forma individual e identificar as características distintas de cada, é necessário delimitar a área de estudo para o desenvolvimento dos MPAs. É importante considerar as características técnicas e operacionais dos trechos rodoviários, assim como, a disponibilidade de dados. Para a realização dos MPAs é necessário que a amostra de dados apresente um tamanho suficiente para o estudo estatístico, isto é, quanto maior a quantidade de eventos ocorridos, maior a probabilidade de se obter variáveis com significância estatística.

Na figura 3.6 podemos observar o mapa da autoestrada A44, respetivas localizações dos sublanços e pontos quilométricos correspondentes. A A44 tem uma extensão de 3,927km, a menos extensa da concessão Costa de Prata, ligando Gulpilhares a Coimbrões. Esta autoestrada em Vila Nova de Gaia é uma autoestrada da região do Grande Porto, que permite a ligação da zona litoral sul do Porto à cidade invicta, representando uma alternativa ao tráfego que passa do centro de Gaia. É constituída por 5 nós de ligação: A29/A44, EN109 (Francelos), Valadares Norte, Madalena e Coimbrões.



Figura 3.6: Mapa da Concessão Costa de Prata, AE A44. Fonte: SIG Ascendi

3.2.2. ANÁLISE DA SINISTRALIDADE DA A44

Seguindo-se a etapa da análise dos dados da área em estudo, extraiu-se a informação que caracteriza as variáveis. As concessões dispõem dum conjunto de dados organizados em diferentes bases de dados dependendo da informação (infraestrutura, tráfego, acidentes, velocidades).

Identificados os dados dos acidentes, realizou-se a análise de sinistralidade da autoestrada em estudo, considerando o período de 2015 a 2020. No gráfico da figura 3.7, estão presentes os números absolutos de acidentes registados, por tipo, no período considerado. O número total de acidentes no período foi de 244, dos quais 102 foram acidentes com vítimas.

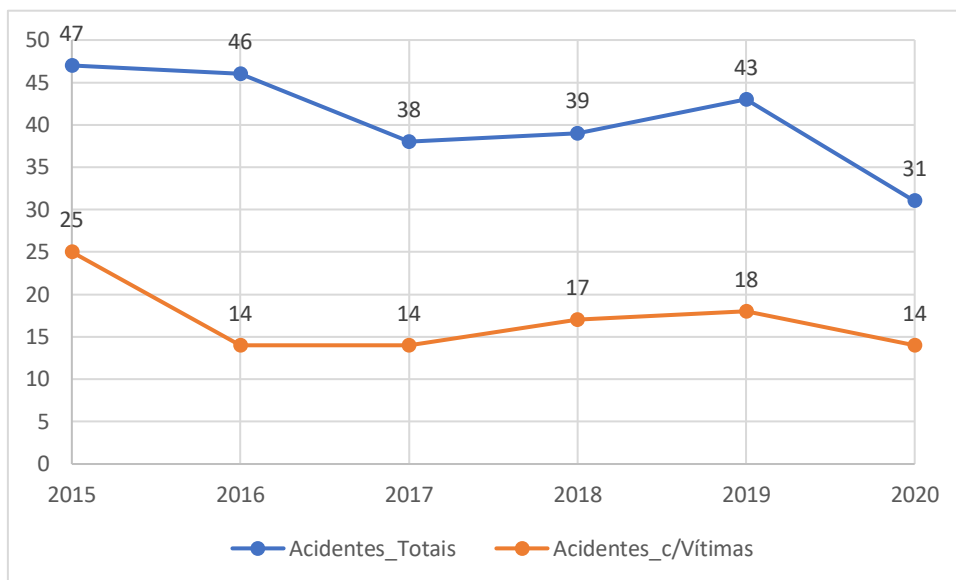


Figura 3.7: Gráfico de sinistralidade na A44 [2015 – 2020]

Observando os resultados, identifica-se que o número total de acidentes, assim como, o número de acidentes com vítimas, apresentam uma tendência decrescente. Porém, é de notar que o número de acidentes com vítimas corresponde a 42,21% do número total de acidentes. É importante referir que, no decorrer do ano 2020, em contexto pandémico, foram aplicadas medidas no combate ao novo coronavírus a nível nacional, resultando numa redução do tráfego médio diário anual (TMDA), contudo, o número de acidentes não demonstrou a redução expectável, comparativamente aos anos antecedentes – figura 3.8.

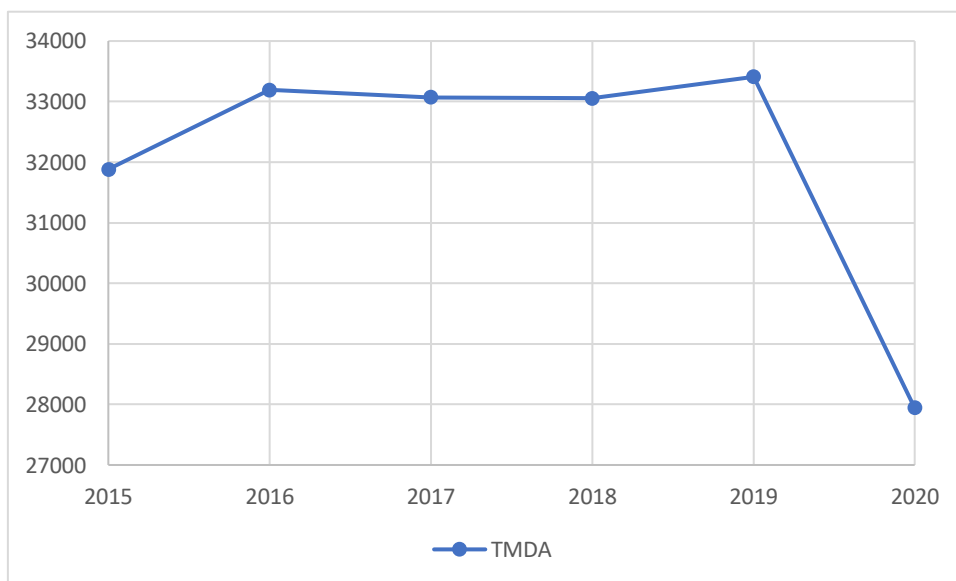


Figura 3.8: Gráfico do TMDA na A44 [2015 – 2020]

É de salientar que o processo de registo dos acidentes é sujeito a erros, quer devido à omissão, quer ao preenchimento incorreto da informação. A falta de informação, por exemplo, da localização do acidente, é um fator prejudicial para a realização do modelo, uma vez que sem a mesma não é possível correlacionar o acidente com a infraestrutura. Quando a localização é registada, é possível obter informação georreferenciada na plataforma SIG, utilizada pela Ascendi. Nesta plataforma são, por sua vez, disponibilizados os relatórios de acidentes, onde estão registadas as características da infraestrutura e as informações que levaram à ocorrência do mesmo.

Na figura 3.9, encontram-se a localização dos acidentes da A44 no período em estudo, segundo a plataforma SIG da Ascendi.



Figura 3.9: Localização dos acidentes rodoviários (Verde: acidentes s/ vítimas, Vermelho: acidentes c/ vítimas). Fonte: SIG Ascendi

3.2.2. RECOLHA DOS DADOS

A partir dos dados em bruto disponíveis em bases de dados distintas, é necessário obter informação da geométrica e operação rodoviária disponível que apresentam relevância para aplicar no modelo. Como tal, foi necessário identificar as variáveis necessárias para o desenvolvimento do modelo e, posterior segmentação da base de dados. As variáveis independentes correspondem à informação sobre a infraestrutura da rede rodoviária, do tráfego e dos limites de velocidade.

A análise dos dados permitiu detetar limitações, de forma a contribuir para a criação de um sistema unificado de dados de operação.

Tráfego

O volume de tráfego é uma variável de exposição fundamental, uma vez que sem a circulação de veículos não ocorrem acidentes. Esta tem sido estudada em diversos trabalhos de modelação, como a mais determinante para a ocorrência de acidentes, sendo o TMDA a forma mais utilizada.

Analizando os dados disponibilizados pela Ascendi, constatou-se que a informação referente ao TMDA está agregada por sublanço, dependendo do sentido e categoria dos veículos (Figura 3.10). A variável “percentagem de pesados” será uma das variáveis explicativas a considerar na modelação.

Localização	Sentido	Categoria Veículo\Ano	2015	2016	2017	2018	2019	2020
Madalena - Coimbrões	Crescente	Ligeiros	40959	43044	42532	43097	39519	33030
		% Pesados	5%	5%	6%	5%	5%	6%
		Total	43040	45208	45012	45254	41658	35136
	Decrescente	Ligeiros	37024	38267	38667	38369	42708	35221
		% Pesados	5%	5%	5%	5%	5%	6%
		Total	38948	40226	40584	40447	44903	37458
Valadares Norte - Madalena	Crescente	Ligeiros	33908	34946	35116	35070	35393	28961
		% Pesados	5%	5%	5%	5%	5%	7%
		Total	35703	36821	37001	36952	37293	31010
	Decrescente	Ligeiros	32686	33699	33871	33826	34136	28692
		% Pesados	6%	6%	6%	6%	6%	6%
		Total	34776	35892	36074	36023	36354	30640
Francelos (EN109) - Valadares Norte	Crescente	Ligeiros	28762	29875	29883	29928	27834	23127
		% Pesados	6%	6%	6%	6%	7%	7%
		Total	30740	31895	31849	31878	30082	24993
	Decrescente	Ligeiros	26748	27650	27700	27394	29893	24833
		% Pesados	7%	7%	7%	7%	6%	7%
		Total	28631	29679	29879	29607	31922	26735
(A29/A44) - Francelos (EN109)	Crescente	Ligeiros	20005	21222	20443	20506	20870	17207
		% Pesados	9%	9%	9%	9%	9%	9%
		Total	21877	23202	22355	22422	22818	18847
	Decrescente	Ligeiros	19545	20738	19976	20039	20393	17146
		% Pesados	8%	8%	8%	8%	8%	8%
		Total	21318	22616	21788	21856	22240	18708

Figura 3.10: Tráfego Médio Diário Anual por sublanço e sentido [2015 – 2020]. Fonte: Ascendi

No gráfico da figura 3.11 podemos observar que o tráfego médio diário anual de pesados apresenta um valor baixo comparativamente ao tráfego médio diário anual total.

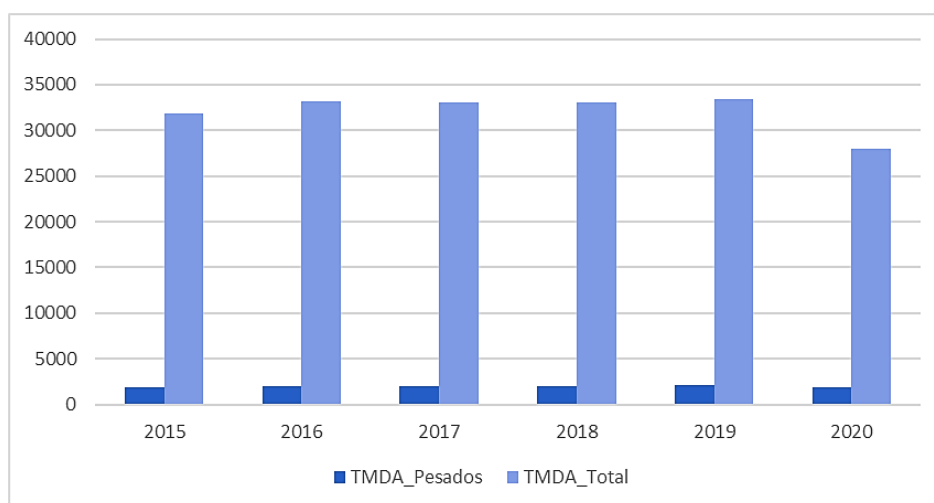


Figura 3.11: Gráfico do Tráfego Médio Diário Anual na A44 [2015 – 2020]

Infraestrutura

Para relacionar os acidentes com as características físicas e funcionais das vias, consideraram-se as características geométricas, o tipo de pavimento e características operacionais. Estas estão associadas por tipo de localização, nós ou arcos (sublanços), sendo este último o foco do presente estudo. Analisando os dados disponíveis, observa-se que os sublanços localizam-se entre nós e correspondem a troços de plena via. A autoestrada A44 é constituída por 4 sublanços que estão apresentados no quadro 3.2.

Quadro 3.2: Sublanços da A44

Tipo_Localizacao	Localizacao	PK_Inicial	PK_Final
Sublanço	A29/A44 - Ligação à EN109 (Francelos)	0	741
Sublanço	Ligação à EN109 (Francelos) - Valadares Norte	741	2053
Sublanço	Valadares Norte - Madalena	2053	3103
Sublanço	Madalena - Coimbrões	3103	3926

Para além das variáveis independentes disponibilizadas, também é habitual considerar como variável de exposição o comprimento dos sublanços, utilizado no desenvolvimento dos modelos.

Para a construção da base de dados foram consideradas as seguintes variáveis independentes, relativas à infraestrutura, presentes nos dados disponíveis:

- Caracterização em planta: reta, curva, clóide, raio e respetivo valor;
- Tipo de pavimento: flexível ou rígido;
- Caracterização de trecho: plena via, garrafão de entrada, garrafão de saída ou área de portagem;
- Caracterização da zona: viaduto/obra de arte, aterro, escavação ou mista;
- Características geométricas e operacionais: tipo de via, número de vias, largura da via, tipo de separador central, largura do separador, tipo de bermas, largura das bermas e velocidade limite permitida.

3.3. SEGMENTAÇÃO DAS VIAS

Este passo é essencial quando se trata de modelação de frequência de acidentes e, por isso, foi necessário adotar uma estratégia de divisão dos trechos, associando os dados a segmentos homogêneos, com extensão variável, tendo em conta as características geométricas das vias e tráfego. Segundo Resende & Benekohal (1997), quanto maior o número de variáveis utilizadas na homogeneidade, menor serão os segmentos homogêneos resultantes. Deste modo, ocorre uma redução da robustez estatística dos MPAs. Além disso, um número maior de segmentos resultará em mais segmentos com zero acidentes, devido à característica de aleatoriedade dos acidentes (Souleyrette et al., 2007).

Tendo por base o *Highway Safety Manual* (HSM) e os dados disponíveis, foi necessário escolher os critérios para homogeneizar e seccionar os segmentos. Dado que será efetuada a divisão da rede rodoviária em segmentos com características homogêneas, deve ter-se em conta o comprimento dos segmentos como variável explicativa para o modelo. É de referir que o HSM refere que a frequência de acidentes é diretamente proporcional ao comprimento do segmento (AASHTO, 2010).

Esta etapa do processo teve o apoio do departamento de informática da empresa que permitiu, de uma forma mais ágil e automática, tratar as bases de dados disponíveis.

A segmentação baseia-se na identificação dos diferentes PK's de início e fim de cada segmento e as suas características determinam a criação um novo segmento, ou seja, quando uma das seguintes características altera, passamos para um segmento novo:

- Caracterização em planta: reta, curva ou clotóide, e raios associados;
- Largura da via;
- Tipo de via;
- Largura das bermas;
- Tipo de berma;
- Largura de separador.

Após a fase de segmentação é necessário identificar as características de trecho (plena via, garrafão de entrada, garrafão de saída, área de portagem), de zona (viaduto/obra de arte, aterro, escavação ou mista) e tipo de pavimento (flexível ou rígido). Para os segmentos resultantes serão acrescentadas as características operacionais para cada segmento: tipo de separador, TMDA, percentagem de pesados, limite de velocidade e presença de equipamentos de controlo eletrónico. Por último, introduz-se os dados referentes à variável dependente necessária para o processo de modelação, os acidentes totais e acidentes com vítimas, correspondentes a cada intervalo de PK.

Esta segmentação deu origem a uma base de dados com 48 segmentos no sentido crescente e 51 no sentido decrescente, resultando num total de 99 segmentos.

Este processo é concluído com a caracterização dos segmentos, para os vários anos analisados, dependendo do sentido da via. Isto é possível, uma vez que a geometria da autoestrada não sofreu alteração. Contudo, é importante avaliar, sempre que necessário, se a geometria e/ou as condições de operação sofreram alteração no ano vigente de análise.

Assim, procedeu-se à análise do período de 2015 a 2020, sendo necessário incluir os seguintes dados para cada segmento resultante:

- TMDA;
- Percentagem de tráfego de veículos pesados;
- Limite de Velocidade;
- Acidentes Totais;
- Acidentes com Vítimas;

Concluído o processo de construção da base de dados, obtiveram-se 594 segmentos para a autoestrada A44. O período em estudo apresenta um total de 244 acidentes totais, sendo o número máximo de acidentes totais, por segmento, igual a 5. Para os acidentes com vítimas a base de

dados contabiliza um total de 102, com um número máximo de 4. Em ambos os casos, o número mínimo de acidentes é zero.

3.4. DADOS CONSIDERADOS NA MODELAÇÃO

Dado que, nos modelos de previsão de frequência de acidentes, os dados são agrupados por segmentos rodoviários, nem os acidentes, nem as suas características, configuram um dado individual. Assim, as características individuais (usuários, veículos e condições climáticas), como já referido, não serão utilizadas como variáveis. Kockelman e Kweon (2002) destacam que a necessidade de agregar os dados por segmentos, tem como consequência, a perda de significância no modelo, das características individuais dos acidentes, dada a impossibilidade de se incorporarem nos atributos dos segmentos.

Desta forma, as variáveis que apresentam importância na aplicação são:

a) Características em planta:

- Reta, curva ou clotóide;
- Comprimento do segmento;
- Raio de curvatura.

É importante considerar o comprimento do segmento, uma vez que é expectável que quanto maior o comprimento maior o número de acidentes registados. Por sua vez, a característica do alinhamento representa um bom indicador de sinistralidade, dada a ligação de curvas a elevadas taxas de acidentes. Ainda, na fase de projeto, existem parâmetros a cumprir referentes à relação entre o raio e a velocidade operacional da via, tornando estas variáveis relevantes.

b) Características do perfil transversal:

- Número de vias;
- Largura separador central;
- Tipo de separador central;
- Largura das bermas;
- Largura da via;
- Outras vias.

Glennon (1987) realizou um estudo, que aponta para a potencial influência dos elementos da seção transversal da via, na ocorrência de acidentes. Tais elementos podem afetar a visão transversal do condutor, conforto e percepção de risco. O número de vias é identificado como uma das variáveis

que está relacionada com o conforto e manobras realizadas na via. Note-se que a variável “tipo de separador” é considerada, na medida em que pode influenciar por exemplo, a velocidade assumida pelo condutor e terá, impacto na gravidade do acidente. Ainda, Odgen (1996) afirma que a berma é caracterizada como uma área que possibilita a recuperação dos veículos que perdem o controlo e garante o afastamento em relação aos objetos e estruturas existentes nas laterais da via. A diminuição da largura da berma aumenta o perigo de acidentes por colisão com objeto fixos (Roque, C. e Mohammad, J., 2018). Por sua vez, o aumento da largura da berma tem impacto na diminuição da velocidade.

c) Características operacionais:

- Tráfego médio diário anual (TMDA)
- Percentagem de pesados
- Limite de velocidade regulamentada

Estas variáveis, associadas às condições de operação das vias, têm um peso considerável na segurança rodoviária, uma vez que influenciam o comportamento dos condutores e como o veículo se desloca. Este tópico tem destaque em alguns estudos, dos quais, Chang e Wang (2006), Das e Abdel-Aty (2010), Das e Abdel-Aty (2011), sendo indicado como um dos fatores de risco de acidentes.

A variável tráfego apresenta uma importante contribuição na definição de MPA e foi calculada ao longo dos quatro sublanços em cada sentido, sendo utilizada para caracterizar, em termos médios, o tráfego nos segmentos. O mesmo se aplica para a variável do tráfego de pesados.

Como é de conhecimento, a velocidade está associada à taxa de sinistralidade e à sua gravidade. Elvik (2009) estudou a relação da velocidade com a segurança rodoviária e concluiu que esta destaca-se como um fator de risco significativa, e o mais relevante para a ocorrência de acidentes, assim como, na gravidade dos mesmos. Neste trabalho, uma vez que não existem registos das velocidades de circulação, será aplicado o limite de velocidade referente a cada sublanço, em km/h.

No quadro 3.3 estão presentes as variáveis aplicadas no processo de modelação:

Quadro 3.3: Descrição das variáveis

Variáveis dependentes	Descrição	Natureza
ACID_TOT	número de acidentes totais	numérica
ACID_VIT	número de acidentes com vítimas	numérica
Variáveis independentes	Descrição	Natureza
LN_TMDA	tráfego médio diário anual	numérica
LN_PERC_PES	percentagem de tráfego de pesados	numérica
LN_COMPR	extensão do segmento	numérica
CARACT_PLANTA	identificação do traçado horizontal do segmento	categórica
RAIO	valor do raio das curvas horizontas	categórica
LARG_BINT	largura da berma interior no segmento	categórica
LARG_BEXT	largura da berma exterior no segmento	categórica
N_VIAS	número de vias no segmento	categórica
LARG_VIA	largura da via	categórica
OUTRAS_VIAS	existência de vias adicionais no segmento	categórica
TIPO_SEP	tipo de separador central no segmento	categórica
LARG_SEP	largura do separador central	categórica
LIMIT_VELOC	limite de velocidade no sublanço	categórica

Os modelos a aplicar baseiam-se na relação entre a variável de resposta e as variáveis independentes. Estas últimas são caracterizadas em dois tipos: variáveis contínuas ou categóricas. No caso do presente estudo, todas as variáveis com natureza categórica, são representadas a partir de variáveis *dummys*, que adotam os valores de 0 a 1. Deste modo, quando a variável categórica apresenta duas características como, por exemplo, a variável “tipo de separador”, que se pode caracterizar como flexível ou rígido, toma valores de 0 para flexível, ou 1 para o rígido.

Por outro lado, quando a variável categórica apresenta n categorias, em que n é maior que 2, são criadas $n-1$ variáveis *dummys* para a representar – Quadro 3.4.

No quadro 3.4 podemos observar a legenda das categorias consideradas no trabalho:

Quadro 3.4: Legenda de categorias

Categoria	0	1	2	3	4	5	6
Sentido	crescente	decrecente					
Característica_Planta		reta	clotóide	curva			
Raio		infinito	até 500	até 1000	até 2000	> 2000	
Tipo de Separador	flexível	rígido					
Largura Berma Exterior		1	1,5	2,5			
Largura Berma Interior	0,7	1					
Outras Vias		nenhuma via a mais	só aceleração	só desaceleração	só lentos	só coletora	=+2 vias
Nº Vias		1	2	3			
Ano		2015	2016	2017	2018	2019	2020
Limite Velocidade		80	90	100	120		

No quadro 3.5 é apresentada a descrição estatística das variáveis independentes utilizadas na modelação.

Quadro 3.5: Descrição estatísticas das variáveis independentes

Variáveis Numéricas	Média	Desvio Padrão	Min - Max
TMDA (veic/dia)	32188,24	6810,478	18708 - 45254
Porcentagem de Pesados (%veic pesados/dia)	0,64773	0,012057	0,05 - 0,09
Comprimento (m)	79,31	50,71987	2 - 196

Variáveis Categóricas	% Observações	
Característica em Planta	Reta	21%
	Clotoide	26%
	Curva	53%
Raio (m)	Infinito	21%
	< 500	36%
	< 1000	4%
	< 2000	11%
	> 2000	27%
Tipo Separador	Flexível	81%
	Rígido	19%
Largura Berma Interior (m)	0,7	98%
	1	2%
Largura Berma Exterior (m)	1	2%
	1,5	17%
	2,5	81%
Outras Vias	nenhuma via a mais	34%
	só aceleração	22%
	só desaceleração	22%
	só coletora	12%
	=+2 vias	9%
Número de Vias	1	4%
	2	93%
	3	3%
Limite de Velocidade (km/h)	80	5%
	90	13%
	100	82%

No quadro 3.5 é possível verificar o intervalado reduzido entre o valor mínimo e máximo do TMDA, que representa a homogeneidade funcional de uma autoestrada. Através dos valores mínimos e máximos da variável *Comprimento*, podemos observar uma heterogeneidade da extensão dos segmentos, que resultaram do processo de segmentação da rede rodoviária.

Da análise geral das variáveis categóricas, utilizadas na modelação da A44, verifica-se que os valores mais elevados de observações por categoria, estão centrados em: curva; raio (<500); separador flexível; largura berma interior igual a 0,7m; largura berma exterior igual a 2,5m; inexistência de outras vias; 2 vias em plena via; limite de velocidade igual a 100km/h.

3.4.1. CORRELAÇÃO ENTRE VARIÁVEIS

Reurings, M., *et al.* (2005) defendem que a escolha das variáveis explicativas deve ter como base um estudo prévio, para que possam ser selecionadas e obtidas de forma fiável e sem que haja correlação com outra variável já incluída.

Com o objetivo de identificar a existência de interdependência entre as variáveis e estimar a influência das mesmas, no processo de modelação, realizou-se uma análise das correlações entre as variáveis explicativas, resultando uma matriz de correlações. Na matriz presente no anexo A, pode-se observar os coeficientes de correlação resultantes [A análise será realizada tendo em consideração valores limite, intervalo de valores entre 0,6 e 0,8 e valores acima de 0,8, sendo que, quando equivalente a 1, as variáveis denominam-se colineares].

Analizando a matriz de correlações, conclui-se que, grande parte dos casos que apresentam valores elevados de correlação, caracterizam o mesmo elemento da via, como é o caso das características em planta e o raio. Pode-se constatar que a variável do *TMDA* é colinear com a variável *percentagem de pesados*, o que significa que estas, quando avaliadas em conjunto, não apresentam significância, sendo necessário optar apenas pela introdução de uma delas. Neste caso, dada a importância da variável *TMDA* no modelo, priorizou-se uma em relação à outra. No que toca às características em planta e raio, em primeiro lugar analisou-se isoladamente as duas variáveis no modelo, tendo-se optado por manter o raio como variável, dada a sequência do estudo realizado por Souza, S. (2021). Esta variável foi selecionada, com o objetivo de permitir o estudo comparativo entre modelos distintos de autoestradas.

A autoestrada A44, comparativamente à A29, analisada anteriormente em Souza, S. (2021), apresenta poucos segmentos, o que poderá dificultar a obtenção de significância estatística na estimativa dos coeficientes das variáveis.

3.5. MODELAÇÃO

Após o tratamento e análise das variáveis, inicia-se o processo de modelação. Os modelos utilizados para o estudo foram o Poisson e BN e têm em vista a elaboração de um modelo de previsão de frequência de acidentes totais e previsão de frequência de acidentes com vítimas.

O processo de modelação será realizado através do *software* NLOGIT 5, onde será introduzida a base de dados, desenvolvida no *Excel*. Nesta etapa do processo é necessário ter atenção à formatação dos dados provenientes, uma vez que podem ocorrer erros na transferência dos mesmos. A cronologia do processo passa por:

- Ajuste dos dados/ variáveis e respetiva formatação;
- Realização de uma análise estatísticas das variáveis, de forma a evitar erros;
- Definição da distribuição a utilizar – Poisson ou BN;
- Análise da estimativa dos coeficientes e sua significância estatística (assumindo-se um nível de confiança superior ou igual a 90%);
- Seleção e formulação do modelo final.

Fridstrom, L., *et al.* (1995) sugerem que as variáveis sejam inseridas uma a uma e apenas são incluídas no modelo se a relação com a variável dependente for estatisticamente significativa. É de salientar que algumas características tratadas como variáveis independentes iniciais para a frequência de acidentes totais e com vítimas não farão parte do modelo, uma vez que não apresentam significância estatística, sendo, por isso, excluídas.

Como abordado no capítulo 2, a relação entre as variáveis dependentes e independentes é representada através da seguinte equação:

$$E[Y]=f(TMDA; COMPR; RAO; OUTRAS_VIAS; LIM_VEL) \quad (3.1)$$

Sendo $E[Y]$ o valor esperado da frequência de acidentes, o *TMDA* e *COMPR* representam as variáveis contínuas de exposição e as restantes correspondem às variáveis categóricas.

Com o objetivo de realizar uma aplicação prática de um modelo de previsão de acidentes e realizar uma etapa validação robusta, foram realizados quatro modelos, com dois períodos de observações distintos. Em primeiro lugar, foram desenvolvidos dois modelos relativos ao período total de observações, 6 anos, para as duas variáveis resposta (acidentes totais e com vítimas). Em segundo lugar, foram desenvolvidos dois modelos (acidentes totais e com vítimas) referentes ao período de 4 anos de observações (2015 a 2018), permitindo realizar a estimação para os anos de 2019 e 2020 e estudar o ajuste dos mesmos na etapa de validação.

3.6. VALIDAÇÃO DO MODELO

O procedimento realizado na validação passa pela aplicação de métricas de desempenho e análise gráfica, para verificação da adequação dos modelos, como sugerido por Humphrey *et al.* (2017). Numa primeira fase, procedeu-se à aplicação das medidas de desempenho e, em seguida procedeu-se à análise gráfica dos gráfico CURE em conjunto com os gráficos de dispersão.

A medição dos indicadores de desempenho do modelo é importante para verificar a capacidade do modelo em prever a frequência de acidentes por comparação com os valores observados. Para esta etapa foram selecionadas algumas métricas, que se ajustavam melhor aos dados observados de acidentes: Erro Relativo (ER); Coeficiente de Determinação (R^2); Erro Logarítmico Quadrático Médio (MSLE); Desvio Absoluto Médio (MAD); Erro Percentual Absoluto Médio Simétrico (SMAPE); Erro Quadrático Médio de Previsão (MSPE).

Relativamente à análise gráfica, o objetivo primário da análise de resíduos é a identificação dos casos em que as estimativas do modelo se distanciam muito dos valores observados – diferença entre o número de acidentes observados e os previstos. Esta análise gráfica começou pela ordenação crescente das variáveis de exposição (TMDA e comprimento do segmento), e consequentemente são ordenados os resíduos. Os resíduos acumulados são calculados como a soma dos resíduos de 1 a j . Um bom ajuste significa que os resíduos acumulados oscilam em torno do eixo horizontal e estão contidos entre os limites definidos por $\pm 2\sigma$, terminando próximo do valor nulo.

Para complementar a análise gráfica, foram elaborados gráficos de dispersão, que permitem a comparação direta entre os valores observados e os previstos, por sua vez, com as variáveis TMDA e comprimento do segmento.

No presente estudo, o processo de validação será aplicado a análise da autoestrada A44 mas também da autoestrada A29, na sequência do estudo anterior realizado por Souza S. (2021). Pretende-se assim comparar o desempenho dos dois modelos para duas autoestadas distintas. Por outro lado, será dada uma particular atenção ao ano 2020. Nesse ano, devido ao contexto pandémico, é expectável a existência de uma relação distinta por exemplo, entre o TMDA e a frequência de acidentes. Como tal, é compreensível que exista um pior desempenho do modelo em prever a frequência de acidentes. Na verdade, não havendo histórico de contexto semelhante, este estudo irá aferir a capacidade dos MPAs tradicionais, como é o caso do modelo de Poisson e BN, para contextos disruptivos.

4

ANÁLISE DE RESULTADOS

Neste capítulo são apresentados os aspectos detalhados dos resultados do processo de modelação e validação.

Nos modelos de resposta quantitativa a variável de resposta corresponde à frequência dos acidentes. Esta é desagregada em função das consequências do acidente, neste caso, acidentes totais (acidentes com vítimas e acidentes apenas com danos materiais) e acidentes com vítimas (mortes, feridos graves e feridos ligeiros).

Neste trabalho, foram desenvolvidos quatro modelos, tendo em conta dados temporais distintos. O estudo foi realizado tendo por base uma amostra temporal de seis anos, período de 2015 a 2020 (modelo de seis anos) e o período de 2015 a 2018 (modelo de quatro anos). Cada um destes períodos teve como resultado um modelo de frequência de acidentes totais e acidentes com vítimas.

O processo de modelação foi realizado através do programa NLOGIT 5, que permite estimar os coeficientes das variáveis independentes. Para a modelação foram seleccionadas as distribuições de Poisson e BN. Para se aferir qual a distribuição mais ajustada à base de dados em estudo, foi aplicado e analisado ao teste χ^2 , que assume como hipótese nula, a igualdade entre a variância da variável dependente e a média. Este teste apresenta como valor padrão $\chi^2_{0,05GL=1}$ (3,84), o que indica que, para valores inferiores a 3,84, a variância é igual à média. Neste trabalho será aplicada a distribuição de Poisson, correspondendo ao melhor ajuste aos dados, de acordo com o referido teste χ^2 .

As variáveis independentes a manter na estrutura do modelo tiveram como processo de seleção, os níveis de confiança dos parâmetros estimados. Estas serão rejeitadas para parâmetros com nível de confiança inferior a 90%. Para valores inferiores a 90% é realizada uma análise com o objetivo

de incluir, ou não, a variável. Contudo, caso a sua presença, sem significância estatística (nível de confiança inferior a 90%), não afetar os níveis de confiança das restantes variáveis, esta mantém-se, uma vez que, pode não apresentar significância estatística neste modelo e ser influente noutra base de dados.

É de notar que a autoestrada A44 estabelece ligação ao meio urbano, fator que pode justificar o comportamento distinto por sublaços, ou seja, diferente tipo de comportamento dos usuários. Este facto pode ter influência nas variáveis e nos MPA.

Para etapa de validação selecionou-se um conjunto de medidas de avaliação de desempenho e análises gráficas, como descrito nos capítulos anteriores, com o objetivo de avaliar a qualidade do ajuste, tendo em conta que os modelos desenvolvidos são MLG. Ainda, é realizada uma análise focada na comparação entre os anos de 2019 e 2020, no sentido de estudar o impacto que as medidas restritivas no combate ao covid-19 tiveram nas estradas.

4.1. MODELO DE PREVISÃO DE FREQUÊNCIA DE ACIDENTES TOTAIS – A44 (15-20)

4.1.1. ETAPA DE MODELAÇÃO

Em primeiro lugar foi desenvolvido o modelo de previsão de frequência de acidentes totais, relativos ao período de 6 anos, com a estrutura do modelo a aplicar, com base na equação 2.9.

O modelo foi aplicado às distribuições de Poisson e BN, verificando-se melhor ajuste para a primeira, após o teste χ^2 (Anexo B.1). A execução do modelo foi ajustada a cada variável independente.

As variáveis independentes, ANO, CARACTERISTICA_PLANTA, LARGURA_BEXT, LARGURA_BINT, N_VIAS, PERC_PESADOS e TIPO_SEP, não apresentaram significância estatística para o modelo (nível de confiança inferior a 90%).

Os resultados desta primeira aplicação (estimativas dos parâmetros do modelo, desvio padrão e respetivos valores P) são apresentados no Quadro 4.1, assim como as medidas de avaliação de desempenho (no Anexo B.1 apresentam-se todos os resultados da aplicação).

Quadro 4.1: Resultados das estimativas do modelo Poisson para acidentes totais

Variável resposta: ACIDENTES TOTAIS			
Váriáveis explicativas	Parâmetro estimado	Desvio padrão	Prob. $ z > Z^*$
Ln α	-19,8978	4,50155	0,0000
TMDA	1,4938	0,43331	0,0006
COMPR	0,83774	0,11832	0,0000
RAIO_C02 (< 500)	-0,36526	0,21405	0,0879
RAIO_C03 (< 1000)	-1,25997	0,39543	0,0014
RAIO_C04 (< 2000)	-0,64394	0,26771	0,0162
RAIO_C05 (> 2000)	-0,64708	0,20773	0,0018
OUTRVI02 (aceleração)	0,51057	0,18863	0,0068
OUTRVI03 (desaceleração)	0,40103	0,18831	0,0332
OUTRVI05 (coletora)	0,02828	0,25193	0,9106
OUTRVI06 (+2 vias)	0,61073	0,27241	0,025
LIMIT_VE01 (80km/h)	0,62594	0,28309	0,027
LIMIT_VE02 (90km/h)	-0,87726	0,28131	0,0018
Medidas de avaliação do desempenho			
Função Logaritimica de verosimilhança			-437,58931
Log likelihood function			-528,03936
Mc Fadden pseudo R ²			0,1712941

Os resultados obtidos para o modelo de previsão de frequência de acidentes totais podem ser representados pela seguinte equação final:

$$E[y] = 2,28^{-9} TMDA^{1,494} Compr^{0,838} e^{-0,3653 Raio_{C2} - 1,26 Raio_{C3} - 0,644 Raio_{C4} - 0,647 Raio_{C5} + 0,5106 OutrVi_{C2} + 0,401 OutrVi_{C3} + 0,0283 OutrVi_{C5} + 0,6107 OutrVi_{C6} + 0,626 Limit_{Ve01} - 0,877 Limit_{Ve02}}$$
(4.1)

sendo:

- $E[y]$ - valor estimado de acidentes totais, por ano;
- $TMDA$ – tráfego médio diário anual do segmento;
- $Compr$ – comprimento do segmento;
- $Raio$ - categoria do raio do segmento (< 500, < 1000, <2000, > 2000), sendo a categoria de referência segmento de reta;
- $Outras_Vias$ – categoria da “outra via” no segmento (aceleração, desaceleração, coletora, +2vias), sendo a categoria de referência o segmento sem outra via;
- $Limit_Veloc$ – limite de velocidade de circulação do segmento (80km/h, 90km/h), sendo a categoria de referência o segmento com limite de velocidade 100km/h;

Na interpretação é importante analisar se os valores estimados para os parâmetros apontam para uma relação das variáveis com os acidentes. Estes parâmetros quando apresentam sinal positivo indicam que a variável independente tem influência no aumento da probabilidade de ocorrer acidentes. Em contrapartida, os parâmetros com valor negativo dizem respeito às variáveis que apresentam um efeito na diminuição dos acidentes. Relativamente às variáveis contínuas, as funções logarítmicas correspondem ao valor estimado do parâmetro, como é o caso do tráfego – TMDA e do COMPR, ou seja, para cada unidade de aumento destas variáveis de exposição, ocorrerá um aumento do valor igual ao coeficiente estimado no logaritmo do valor média da variável dependente, caso as outras variáveis se mantenham constantes.

Com o objetivo de comparar os resultados dos parâmetros estimados, que apresentam estruturas diferentes, foi calculada a elasticidade das variáveis independentes. Deste modo, é possível representar o impacto que as variáveis explicativas apresentam na variável dependente.

Quadro 4.2: Elasticidade das variáveis contínuas - Acidentes Totais

Variável	Elasticidade
TMDA	1,4938
COMPR	0,83774

Considerando o valor da elasticidade do TMDA, pode-se concluir que, um aumento de 10% desta variável, corresponde um aumento de 15% do número de acidentes. Este valor, maior que 1, traduz uma relação não linear entre os acidentes e o TMDA.

O valor da elasticidade da variável COMPR indica que, a um aumento de 10% desta variável, corresponde um aumento de 8,4% do número de acidentes. Dado que o seu valor de elasticidade se encontra próximo de 1, revela uma relação quase linear entre os acidentes e o comprimento do segmento. Este resultado é coerente com o apresentado por Mountain, L., et al. (1996).

Relativamente às variáveis *dummys* é calculada a pseudo-elasticidade, com base na equação 2.14 e cujos valores estão apresentados no quadro 4.3:

Quadro 4.3: Pseudo-elasticidade das variáveis dummies - Acidentes Totais

Variável	Pseudo-Elasticidade
RAIO_C02 (< 500)	-0,4409
RAIO_C03 (< 1000)	-2,5253
RAIO_C04 (< 2000)	-0,9040
RAIO_C05 (> 2000)	-0,9100
OUTRVI02 (aceleração)	0,3998
OUTRVI03 (desaceleração)	0,3304
OUTRVI05 (coletora)	0,0279
OUTRVI06 (+2 vias)	0,4570
LIMIT_VE01 (80km/h)	0,4652
LIMIT_VE02 (90km/h)	-1,4043

Para este modelo, a variável ANO não demonstrou significância estatística, tendo sido, por isso, excluída do modelo. Este resultado poderá revelar que não há tendência, quer de diminuição, quer de aumento dos acidentes, no período analisado, resultante da estabilidade dos valores observados dos acidentes, que não apresentam flutuação ao longo dos diferentes anos (Figura 3.7).

Realizando uma análise detalhada para cada variável, podemos constatar que o valor da elasticidade, neste caso, designada por pseudo-elasticidade, aponta quer para um aumento, quer para decréscimo dos acidentes, mantendo constante todas as outras variáveis

No que diz respeito à variável RAIO, os valores foram desagregados por categoria de raio presente na autoestrada, considerando o raio infinito (ou seja, segmento em reta) como valor de referência. Os valores de referência são os dados selecionados como padrão de comparação, uma vez que são os valores predominantes na autoestrada. Pode-se constatar que valores de raio <500 até raio >2000 reproduzem um decréscimo dos acidentes, em relação a segmentos com raio infinito. Analisando a magnitude deste efeito pelos valores presentes no quadro 4.3, verifica-se que para RAIO_C02 (<500), o decréscimo nos acidentes é de 44% em relação a segmentos com raio infinito. Para valores de RAIO_C03 há uma diminuição nos acidentes acima de 200% (253%). Já para os RAIO_C04 e RAIO_C05 verifica-se uma pseudo-elasticidade de -0,904 e -0,91, respetivamente, o que revela uma diminuição na ordem dos 90% dos acidentes. Estes resultados vão contra a expectativa do efeito das curvas nos acidentes, pois a maior parte dos estudos apontam para o efeito contrário (Souza, S. 2021; Cardoso, J. 2009). Este efeito poderá advir do facto de os condutores apresentarem uma condução preventiva, quando circulam em curva e em clotóide, sendo que, em reta poderão assumir comportamentos de risco, tais como, aumento de velocidade.

Em contrapartida, para as variáveis OUTRAS_VIAS, pode-se comprovar que as várias categorias desagregadas, considerando como referência “nenhuma via a mais”, representam um acréscimo

nos acidentes, de acordo com o esperado. Estão presentes quatro categorias, OUTFVI02 (aceleração), OUTFVI03 (desaceleração), OUTFVI05 (coletora) e OUTFVI06 (+2 vias), correspondendo a valores percentuais de pseudo-elasticidades de, 40%, 33%, 2,8% e 46%, respectivamente. Esta variável tem influência nos acidentes, dada a transição de perfil transversal em relação à plena via e com esta, um aumento potencial do número de conflitos entre os veículos em mudança de via.

Por último, para a variável LIMIT_VELOC, pode constatar-se dois efeitos distintos nos acidentes, tendo como referência vias com limite de velocidade de 100km/h. Por um lado, o LIMIT_VE01, correspondente a 80km/h, representa um aumento de 47% nos acidentes, e, por outro lado, o LIMIT_VE02 (ou seja, 90km/h) representa uma diminuição de 140% dos acidentes, e pode estar localizado em zonas com características geométricas menos perigosas (geometria mais “correta”). Estes efeitos distintos na sinistralidade da diminuição do limite de velocidade, relativamente ao valor 100km/h, necessitam de uma análise mais profunda, nomeadamente, com visitas ao local. Geralmente, os limites de velocidade são definidos de acordo com as características geométricas do local, grande parte delas, representadas no modelo em estudo. Seria ainda relevante, realizar medições de velocidades praticadas que poderão indiciar um desajuste destas, em relação aos limites de velocidade estabelecidos.

4.1.2. ETAPA DE AVALIAÇÃO DO DESEMPENHO DO MODELO

Nesta etapa de avaliação procedeu-se à análise de resíduos com o objetivo de aferir a qualidade de ajuste do modelo aos dados. Este método consiste na elaboração gráfica de resíduos acumulados dos acidentes, ordenados, neste caso, de acordo com o TMDA e o COMPR, figura 4.1 e figura 4.2, respetivamente.

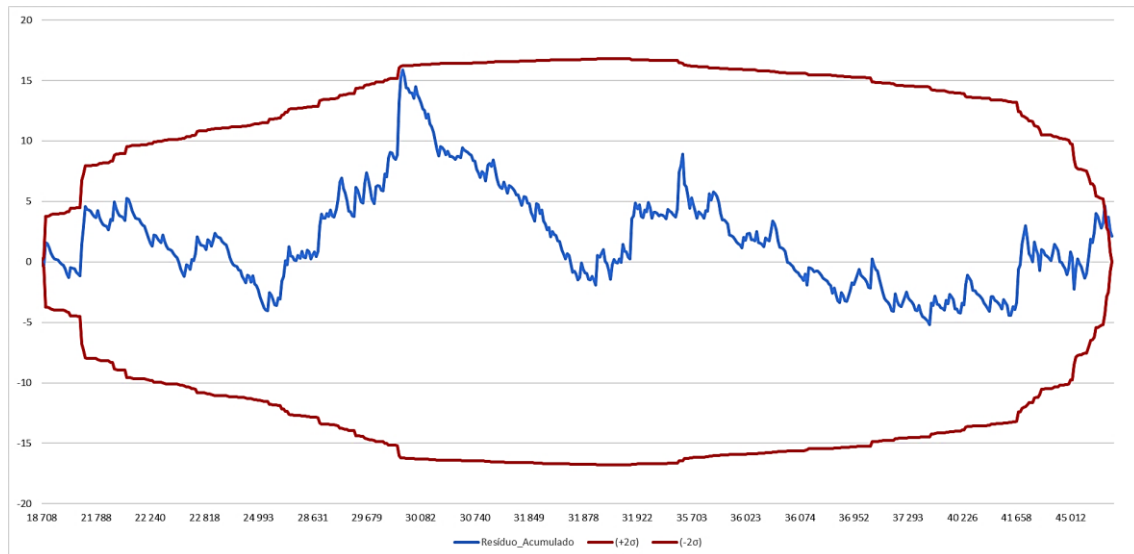


Figura 4.1: - Gráfico resíduos acumulados da variável TMDA para acidentes totais

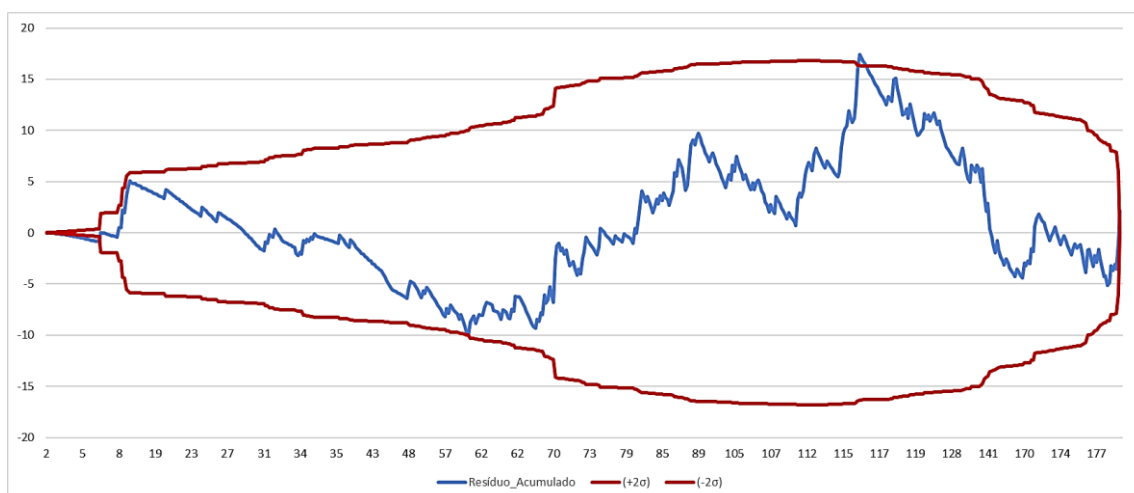


Figura 4.2: - Gráfico resíduos acumulados da variável comprimento para acidentes totais

O gráfico de resíduos acumulados, em ordem ao TMDA, apresenta uma curva segundo o padrão desejado, ou seja, os resíduos acumulados encontram-se distribuídos em torno do eixo horizontal e estão contidos entre as bandas definidas por $\pm 2\sigma$. É possível observar que, para valores de tráfego próximos de 30 000, os resíduos acumulados aproximam-se da banda delimitadora $+2\sigma$, no entanto, pode-se concluir que o modelo apresenta, de forma geral, um bom ajuste.

Relativamente à variável comprimento, observa-se uma distribuição, para alguns valores, afastada do eixo do zero e próxima dos limites definidos por $\pm 2\sigma$. Estes dizem respeito ao intervalo de comprimentos entre 45 e 70m de segmento, e entre 110 e 120m. Contudo, pode-se considerar que, de uma forma geral, o modelo apresenta um bom desempenho.

4.2. MODELO DE PREVISÃO DE FREQUÊNCIA DE ACIDENTES COM VÍTIMAS – A44 (15-20)

4.2.1. ETAPA DE MODELAÇÃO

Neste capítulo foi desenvolvido o modelo de previsão de frequência de acidentes com vítimas, relativos ao período de 6 anos, mantendo a mesma estrutura do modelo anterior.

O modelo foi aplicado às distribuições de Poisson e BN, verificando-se melhor ajuste para a primeira, após o teste χ^2 (Anexo B.2). A execução do modelo foi ajustada a cada variável explicativa.

A par com o modelo desenvolvido anteriormente, as variáveis explicativas, ANO, CARACTERISTICA_PLANTA, LARGURA_BEXT, LARGURA_BINT, N_VIAS, PERC_PESADOS, mantêm-se sem significância estatística. Contudo a variável TIPO_SEP passa a apresentar significância para o modelo (nível de confiança superior a 90%), ao contrário da variável OUTRAS_VIAS, que deixa de apresentar significância neste modelo.

Os resultados desta primeira aplicação (estimativas dos parâmetros do modelo, desvio padrão e respetivos valores P) são apresentados no Quadro 4.4, assim como as medidas de avaliação de desempenho (no Anexo B.2 apresentam-se todos os resultados da aplicação).

Quadro 4.4: Resultados das estimativas do modelo Poisson para acidentes com vítimas

Variável resposta: ACIDENTES COM VÍTIMAS				
Váriáveis explicativas	Parâmetro estimado	Desvio padrão	Prob. $ z > Z^*$	
Ln α	-27,1364	5,60506	0,0000	
TMDA	2,23057	0,52486	0,0000	
COMPR	0,7456	0,16874	0,0000	
TIPO_SEP	-0,6111	0,28502	0,0320	
RAIO_C02 (< 500)	-0,97853	0,29396	0,0009	
RAIO_C03 (< 1000)	-1,93441	0,60747	0,0015	
RAIO_C04 (< 2000)	-1,25452	0,42400	0,0031	
RAIO_C05 (> 2000)	-1,13621	0,26805	0,0000	
LIMIT_VE02 (90km/h)	-1,55247	0,43363	0,0003	
Medidas de avaliação do desempenho				
Função Logaritimica de verosimilhança			-250,67689	
Log likelihood function			-295,98287	
Mc Fadden pseudo R^2			0,1530696	

Os resultados obtidos para o modelo de previsão de frequência de acidentes com vítimas podem ser representados pela seguinte equação final:

$$E[y] = 1,64^{-12} TMDA^{2,23} Compr^{0,7456} e^{-0,6111 Tipo_Sep - 0,979 Raio_{C2} - 1,934 Raio_{C3} - 1,255 Raio_{C4} - 1,136 Raio_{C5} - 1,553 Limit_{ve02}} \quad (4.2)$$

Sendo:

- $E[y]$ - valor estimado de acidentes com vítimas, por ano;
- *TMDA* – tráfego médio diário anual do segmento;
- *Compr* – comprimento do segmento;
- *Tipo_Sep* – característica do separador, sendo a categoria de referência separador flexível;
- *Raio* - categoria do raio do segmento (< 500, < 1000, <2000, > 2000), sendo a categoria de referência segmento em reta;
- *Limit_Veloc* – categoria do limite de velocidade de circulação do segmento (80km/h, 90km/h), sendo a categoria de referência o segmento com limite de velocidade de 100 km/h.

Com o objetivo de comparar os resultados dos parâmetros estimados, que apresentam estruturas diferentes, foi calculada a elasticidade das variáveis independentes. Deste modo, é possível representar o impacto que as variáveis explicativas apresentam na variável dependente.

Quadro 4.5: Elasticidade das variáveis contínuas - Acidentes com Vítimas

Variável	Elasticidade
TMDA	2,23057
COMPR	0,7456

Considerando o valor da elasticidade do TMDA, pode-se concluir que, um aumento de 10% desta variável, corresponde um aumento de 22% do número de acidentes com vítimas. Este valor, maior que 1, traduz uma relação não linear entre os acidentes e o TMDA.

O valor da elasticidade da variável COMPR indica que, a um aumento de 10% desta variável, corresponde um aumento de 7,5% do número de acidentes com vítimas. Dado que o seu valor de elasticidade se encontra próximo de 1, revela uma relação quase linear entre os acidentes com vítimas e o comprimento do segmento.

Relativamente às variáveis *dummys* é calculada a pseudo-elasticidade, com base na equação 2.14 e cujos valores estão apresentados no quadro 4.6:

Quadro 4.6: Pseudo-elasticidade das variáveis dummys - Acidentes com Vítimas

Variável	Pseudo-Elasticidade
TIPO_SEP	-0,842456987
RAIO_C02 (< 500)	-1,660542369
RAIO_C03 (< 1000)	-5,919960074
RAIO_C04 (< 2000)	-2,506155016
RAIO_C05 (> 2000)	-2,114940341
LIMIT_VE02 (90km/h)	-3,723121898

Analisando a variável TIPO_SEP verifica-se uma elasticidade de -0,8425, o que indica um decréscimo de 84% dos acidentes com vítimas, para vias com separador rígido, em relação às vias com separador flexível, que se apresentam como referência. O separador rígido na rede em estudo, permite realizar a separação entre a plena via e outras vias. Esta diminuição da sinistralidade na presença de separador rígido, pode dever-se ao facto deste tipo de separador estar localizado em zonas mais urbanas, levando os condutores a adotar um comportamento mais defensivo, tal como, diminuição de velocidade.

No que diz respeito à variável RAIO, os valores descrevem a mesma tendência decrescente nos acidentes com vítimas, comparativamente a segmentos com raio infinito. As suas pseudo-elasticidades apresentam valores bastante elevados e contrários ao expectável. Com o objetivo de interpretar de melhor forma os resultados, realizou-se um estudo visual da distribuição dos acidentes nos sublanços, através do SIG. Desta pesquisa consegue-se perceber uma acumulação de acidentes em alinhamento reto, verificando-se que em curva a ocorrência dos mesmos é diminuta.

Para a variável LIMIT_VELO02 o valor da pseudo-elasticidade corresponde a -3,723, o que representa um decréscimo de 372% de acidentes com vítimas, para vias com velocidade limite de 90km/h em relação a vias com limite de velocidade 100km/h. Para o limite de velocidade de 80km/h não se verificou significância estatística. O decréscimo de acidentes com vítimas, em resultado da diminuição do limite de velocidade, está de acordo com as expectativas e estudos anteriores (Souza, S. 2021, Cardoso, J. 2009).

4.2.2. ETAPA DE AVALIAÇÃO DO DESEMPENHO DO MODELO

Nesta etapa de avaliação procedeu-se à análise de resíduos com o objetivo de aferir a qualidade de ajuste do modelo. Este método consiste na elaboração gráfica de resíduos acumulados dos acidentes, ordenados, neste caso, de acordo com o TMDA e o COMPR, figura 4.3 e figura 4.4, respetivamente.

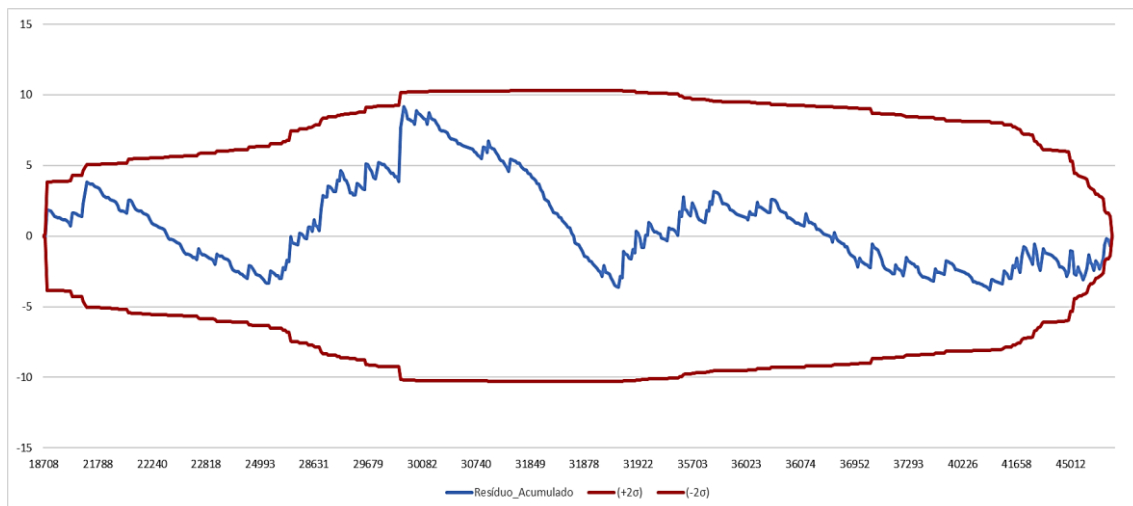


Figura 4.3: - Gráfico resíduos acumulados da variável TMDA para acidentes com vítimas

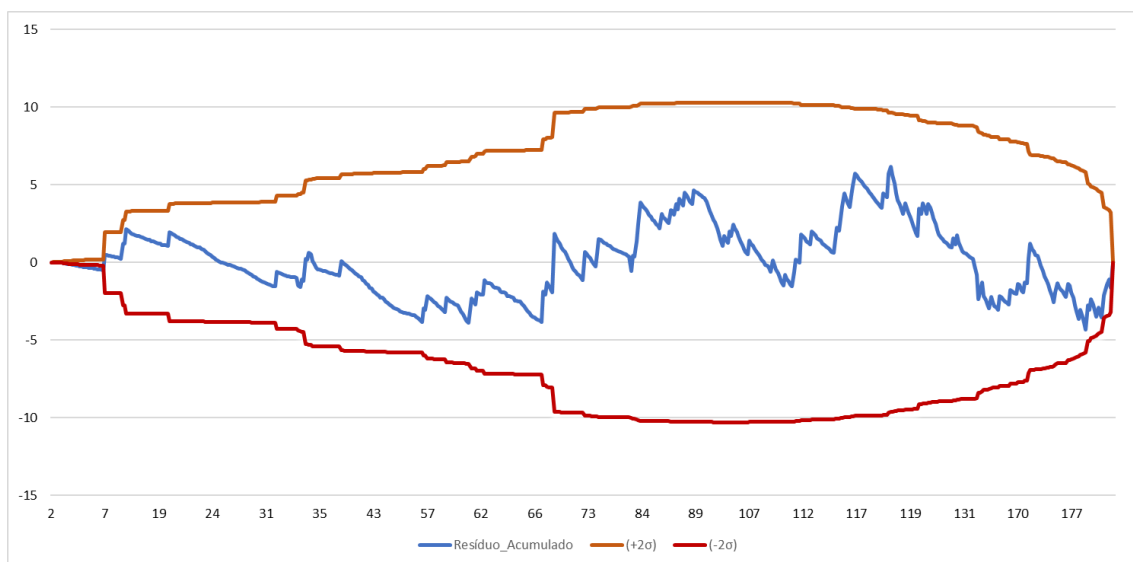


Figura 4.4: - Gráfico resíduos acumulados da variável comprimento para acidentes com vítimas

Realizando uma análise gráfica, podemos comprovar que os resíduos acumulados encontram-se distribuídos em torno do eixo horizontal, respeitando os limites definidos por $\pm 2\sigma$, para as duas variáveis contínuas. Note-se, também, que os valores finais dos resíduos acumulados terminam próximo do valor nulo, comprovando um bom ajuste do modelo aplicado.

4.3. VALIDAÇÃO DOS MODELOS DE PREVISÃO DE ACIDENTES

Neste capítulo é realizado um processo de validação mais robusto, dos modelos de previsão de acidentes, permitindo desta forma, futuras aplicações dos mesmo, como ferramentas de previsão de acidentes.

São realizadas três análises para duas autoestradas distintas. A primeira aplicação prática diz respeito à autoestrada A44, analisada ao longo deste trabalho, e considera duas variáveis resposta, acidentes totais e acidentes com vítimas. Para estes dois modelos foram utilizados os primeiros quatro anos de observações (2015, 2016, 2017, 2018), restando os anos de 2019 e 2020 para estimação dos acidentes. A opção de dividir os 6 anos em 4+2, em alternativa a 3+3, deve-se ao facto de o número de segmentos desta autoestrada ser reduzido, não sendo possível obter estimativas dos parâmetros com significância estatística, na maior parte das variáveis, com base num histórico de 3 anos.

Por último, é realizada a validação dos modelos para a autoestrada A29, desenvolvidos, anteriormente, no trabalho de Souza, S (2021), utilizando os mesmos indicadores de desempenho, podendo, deste modo, avaliar e comparar os modelos das duas autoestradas. Para esta autoestrada será aplicado o modelo de previsão de frequência de acidentes totais, utilizando os primeiros três anos de observações (2015, 2016, 2017), estimando os acidentes para os anos de 2019 e 2020. No caso desta autoestrada, o número de segmentos é maior tendo sido possível separar o período de 6 anos em 3+3, ou seja, 3 anos para estimar o modelo, e 3 anos para previsão dos acidentes, e respetiva validação. Esta separação foi já considerada no trabalho de Souza, S. (2021). É de referir que no presente trabalho, considerou-se apenas a previsão para os anos 2019 e 2020, de forma a ser comparável os resultados das duas autoestradas.

Na etapa de validação são aplicadas medidas de desempenho, assim como, análises gráficas, que permitem validar os modelos, comparando os valores estimados com os observados para estes anos. Ainda que não haja consenso sobre uma definição precisa na literatura (Cronbach, 1971; Pedhazur & Schmelkin, 1991), a validação tem por objetivo apurar a adequabilidade de um modelo ao propósito para o qual foi desenvolvido, fornecendo segurança na sua utilização e interpretação dos resultados.

É de notar, que neste capítulo, o processo de validação é aplicado a dois anos muito distintos – 2019, ano anterior à pandemia covid-19, e 2020 – ano de pandemia. Esta última análise tem como objetivo perceber se o ano atípico de 2020, também em termos de circulação rodoviária, traduz-se numa relação distinta entre o TMDA e a frequência de acidentes, cujos modelos tradicionais não possam representar.

4.3.1. ESTIMAÇÃO DO MODELO DE PREVISÃO DE FREQUÊNCIA DE ACIDENTES

AUTOESTRADA A44

- Acidentes totais - A44 (15-18)**

Para o modelo de previsão de acidentes totais, relativos ao período de 4 anos, pode observar-se no Quadro 4.7, os resultados das estimativas dos parâmetros do modelo. (no Anexo B.3 apresentam-se todos os resultados da aplicação).

Quadro 4.7: Resultados das estimativas do modelo Poisson para acidentes totais – A44

Variável resposta: ACIDENTES TOTAIS				
Váriáveis explicativas	Parâmetro estimado	Desvio padrão	Prob. z >Z*	
Ln α	-17,4057	4,84352	0,0003	
TMDA	1,23663	0,46703	0,0081	
COMPR	0,85254	0,13149	0,0000	
RAIO_C02 (< 500)	-0,3283	0,23566	0,1636	
RAIO_C03 (< 1000)	-1,29438	0,45205	0,0042	
RAIO_C04 (< 2000)	-0,45377	0,28842	0,1157	
RAIO_C05 (> 2000)	-0,58806	0,22626	0,0093	
OUTRVI02 (aceleração)	0,57197	0,20867	0,0061	
OUTRVI03 (desaceleração)	0,49626	0,20867	0,0174	
OUTRVI05 (coletora)	0,00995	0,28448	0,9721	
OUTRVI06 (+2 vias)	0,8022	0,2937	0,0063	
LIMIT_VE01 (80km/h)	0,65442	0,3128	0,0364	
LIMIT_VE02 (90km/h)	-0,74012	0,29923	0,0134	
Medidas de avaliação do desempenho				
Função Logaritmica de verosimilhança			-362,72871	
Log likelihood function			-435,69631	
Mc Fadden pseudo R ²			0,1674735	

Os resultados obtidos para o modelo de previsão de frequência de acidentes totais podem ser representados pela seguinte equação final:

$$E[y] = 2,76^{-8} TMDA^{1,237} Compr^{0,853} e^{-0,328 Raio_{C2} - 1,294 Raio_{C3} - 0,454 Raio_{C4} - 0,588 Raio_{C5} + 0,572 OutrVi_{C2} + 0,496 OutrVi_{C3} + 0,01 OutrVi_{C5} + 0,802 OutrVi_{C6} + 0,654 Limit_{VE01} - 0,74 Limit_{VE02}} \quad (4.3)$$

Através da equação 4.3 foram calculados os valores dos acidentes esperados para cada segmento, e para cada ano em estudo.

- **Acidentes com vítimas – A44 (15-18)**

Para o modelo de previsão de acidentes com vítimas, relativos ao período de 4 anos, pode-se observar no Quadro 4.8, os resultados das estimativas dos parâmetros do modelo (no Anexo B.4 apresentam-se todos os resultados da aplicação).

Quadro 4.8: Resultados das estimativas do modelo Poisson para acidentes com vítimas – A44

Variável resposta: ACIDENTES COM VÍTIMAS				
Váriáveis explicativas	Parâmetro estimado	Desvio padrão	Prob.	z >Z*
Ln α	-24,2411	6,00077	0,0001	
TMDA	1,95599	0,56241	0,0005	
COMPR	0,73189	0,18422	0,0001	
TIPO_SEP	-0,64984	0,31525	0,0393	
RAIO_C02 (< 500)	-1,09446	0,33326	0,0010	
RAIO_C03 (< 1000)	-2,12025	0,73752	0,0040	
RAIO_C04 (< 2000)	-1,01025	0,43490	0,0202	
RAIO_C05 (> 2000)	-0,99307	0,28803	0,0006	
LIMIT_VE02 (90km/h)	-1,33768	0,43668	0,0022	
Medidas de avaliação do desempenho				
Função Logaritimica de verosimilhança			-208,52564	
Log likelihood function			-244,49006	
Mc Fadden pseudo R ²			0,1470997	

Os resultados obtidos para o modelo de previsão de frequência de acidentes com vítimas podem ser representados pela seguinte equação final:

$$E[y] = 2,97^{-11} TMDA^{1,96} Compr^{0,732} e^{-0,65 Tipo_{sep} - 1,095 Raio_{C2} - 2,12 Raio_{C3} - 1,01 Raio_{C4} - 0,993 Raio_{C5} - 1,338 Limit_{ve02}} \quad (4.4)$$

Através da equação 4.4 foram calculados os valores dos acidentes com vítimas esperados para cada segmento, e para cada ano em estudo.

AUTOESTRADA A29

- **Acidentes totais – A29 (15-17)**

Para o modelo de previsão de acidentes totais, relativos ao período de 3 anos, pode-se observar no Quadro 4.9, os resultados das estimativas dos parâmetros do modelo (no Anexo B.5 apresentam-se todos os resultados da aplicação).

Quadro 4.9: Resultados das estimativas do modelo Poisson para acidentes totais – A29

Variável resposta: ACIDENTES TOTAIS				
Váriáveis explicativas	Parâmetro estimado	Desvio padrão	Prob.	z >Z*
Ln α	-13,238	1,990	0,000	
TMDA	0,760	0,197	0,000	
COMPR	0,947	0,065	0,000	
Ln ANO	-0,065	0,031	0,037	
N_VIAS (3 vias)	0,757	0,252	0,003	
LARG_VIA (3,75 m)	-0,932	0,271	0,001	
RAIO (< 500m)	0,238	0,097	0,014	
L_BER (3,75m)	-0,228	0,178	0,201	
VELOC_LIMIT (120km/h)	0,310	0,156	0,048	
Medidas de avaliação do desempenho				
Função Logaritmica de verosimilhança			-1262,70	
Log likelihood function			-1435,53	
Mc Fadden pseudo R ²			0,12	

Os resultados obtidos para o modelo de previsão de frequência de acidentes com vítimas podem ser representados pela seguinte equação final:

$$E[y] = 1,8 \times 10^{-6} TMDA^{0,760} Compr^{0,947} 0,937^{ANO} e^{0,238 Raio - 0,932 LargVia - 0,228 LargBerma + 0,757 NVias + 0,310 Veloc} \quad (4.5)$$

Através da equação 4.5 foram calculados os valores dos acidentes com vítimas esperados para cada segmento, para cada ano em estudo.

4.3.2. VALIDAÇÃO DOS MODELOS DE PREVISÃO DE FREQUÊNCIA DE ACIDENTES

Neste capítulo apresenta-se o processo de validação dos modelos, para a autoestrada A44 e para a autoestrada A29, com base na aplicação de sete medidas de avaliação do desempenho da previsão, e análises gráficas, que comparam os valores previstos com os observados.

No quadro 4.10 estão presentes os valores obtidos para as medidas designadas de coeficiente de determinação (R^2) erro relativo, para os três modelos de previsão desenvolvidos. Estas métricas

estabelecem a relação entre os valores observados e os valores previstos da variável resposta, acidentes totais ou acidentes com vítimas.

Quadro 4.10: Erros relativos e coeficientes de determinação modelos de previsão [2019 – 2020]

	A44_AcidTotais (4anos)		A44_AcidVítimas (4anos)		A29_AcidTotais (3anos)	
	2019	2020	2019	2020	2019	2020
R ²	0,2666	0,2668	0,1415	0,0815	0,0947	0,0748
ER	0,7580	1,0193	0,7655	0,9074	0,9312	0,9396

O coeficiente de determinação avalia a correlação existente entre os dados observados e previstos. Quanto mais próximo de 1, mais ajustado é o modelo. Verifica-se que, de uma forma geral, os valores de R² são reduzidos, ou mesmo próximos de zero, revelando um fraco ajuste dos modelos, com agravamento para o ano 2020.

A métrica ER é um índice relativo em que, quanto mais próximo de zero, mais aceitável é o modelo. Desta forma, podemos comprovar que o parâmetro revela um desajuste para todos os modelos de previsão, com valores mais elevados para o ano de 2020. Estes valores quando próximos de um, demonstram que o modelo apresenta um desempenho quase tão bom, quanto a simples média dos valores observados. É de notar a exceção do modelo de previsão de frequência de acidentes totais, para a autoestrada A44, que no ano 2020, ultrapassa o valor de um, o que se conclui que a média observada é melhor preditor que o próprio modelo.

O valor de ER alto e R² muito baixo reforça a dificuldade de ajuste dos modelos de previsão.

Realizou-se ainda uma análise comparativa entre os valores observados e previstos para os 2 anos com base nos indicadores estatísticos – média, desvio padrão, valor mínimo e máximo.

No quadro 4.11 é possível observar as estatísticas resumidas para os três modelos desenvolvidos.

Quadro 4.11: Estatísticas resumidas dos valores observados e previstos [2019 – 2020]

A44_AcidTotais (4anos)	Valores Observados_2019	Valores Previstos_2019	Valores Observados_2020	Valores Previstos_2020
Média	0,4343	0,4467	0,3232	0,3585
Mínimo	0	0,0221	0	0,0179
Máximo	4	3,0598	4	2,4788
Desvio Padrão	0,8591	0,4434	0,6973	0,3585
A44_AcidVítimas (4anos)				
Média	0,1818	0,1808	0,1515	0,1276
Mínimo	0	0,0059	0	0,0042
Máximo	2	1,2059	2	0,8643
Desvio Padrão	0,4816	0,1812	0,4599	0,1291
A29_AciTotais (3anos)				
Média	0,1614	0,1518	0,1257	0,1196
Mínimo	0	0,0011	0	0,0008
Máximo	3	1,1905	3	0,9374
Desvio Padrão	0,4455	0,1368	0,3948	0,1078

Efetuada uma breve análise dos dados presentes no quadro 4.11, pode verificar-se que os valores são muito próximos, quando se compara os dados observados e previstos. As maiores diferenças estão presentes essencialmente nos valores do desvio padrão, já que nos valores observados, os desvios são maiores, comparativamente aos valores previsto, uma vez que os primeiros refletem a aleatoriedade da ocorrência dos acidentes.

Calculou-se, ainda, outras métricas de erro e comportamento dos resíduos. Nesse sentido, foram obtidas e analisadas medidas de desempenho, apresentadas no quadro 4.12.

Quadro 4.12: Medidas de desempenho modelos de previsão [2019 - 2020]

	A44_AcidTotais (4anos)		A44_AcidVítimas (4anos)		A29_AcidTotais (3anos)	
	2019	2020	2019	2020	2019	2020
MSLE	0,1541	0,1490	0,0690	0,0676	0,0693	0,0564
MAD	0,5032	0,4635	0,2629	0,2295	0,2466	0,2035
SMAPE	1,6525	1,7706	1,8895	1,9386	1,9071	1,9403
MSPE	0,5537	0,4907	0,1757	0,1900	0,1846	0,1462

O MSLE, idealmente, deve ser próximo de zero, o que significa que o valor previsto é próximo do observado. Analisando o quadro pode-se constatar que os modelos apresentam valores mais baixos para o ano de 2020 e mais próximos de zero para os acidentes com vítimas na A44 e acidentes totais na A29.

No caso do MAD (desvio absoluto médio) pode-se comprovar que, quanto maior o MAD, maior a dispersão dos dados. É possível observar que o ano 2019 apresenta valores mais elevados, em especial, para os acidentes totais na A44. Também neste caso, conclui-se que a previsão para o ano 2020 é mais ajustada, apresentando valores mais baixos, para os acidentes com vítimas na A44 e acidentes totais na A29.

O SMAPE é uma medida de precisão baseada em erros relativos. Por sua vez, quanto menor o valor, maior a precisão do modelo. Pelo quadro, verifica-se menores valores para o ano 2019 e para o modelo de acidentes totais na A44, o que difere das conclusões retiradas pelas outras métricas.

Também no caso da medida MSPE, quanto menores os valores desta métrica, melhor a precisão do modelo em prever os acidentes. Neste caso, os modelos de acidentes com vítimas na A44 e acidentes totais na A29 são os que apresentam menores valores. Em suma, não se salienta um ano melhor do que o outro, ou seja, altera com a variável resposta.

Realizando uma análise conjunta das diferenças métricas, destaca-se o predomínio dos melhores resultados, mesmo que ligeiros, para o ano 2020. Apesar das alterações verificadas na mobilidade para este ano, os modelos conseguiram estabelecer a mesma fiabilidade, comparativamente a um

ano comum. Verifica-se, ainda, que os modelos de acidentes com vítimas na A44 e acidentes totais na A29 são, em geral, mais eficientes a prever, do que o modelo de acidentes totais na A44. É de salientar que as métricas ER e R^2 indicam uma aparente ineficiência do modelo, em prever os acidentes, contrariando as outras métricas, incluindo a estatística resumida.

Modelo de Previsão de Acidentes Totais – A44 (15-18)

Após as constatações iniciais, segue-se a análise de resíduos, assim como, a elaboração de gráficos de dispersão para os valores de acidentes observados e previstos, para cada modelo desenvolvido. Através destes é possível analisar os resíduos acumulados em função da variável de interesse, neste trabalho, o TMDA e o COMPR, sendo esperada uma oscilação em torno de zero e dentro dos limites estabelecidos.

Nas figuras seguintes são apresentados os gráficos de resíduos acumulados para cada variável de exposição, associado ao ano de 2019 para a autoestrada A44.

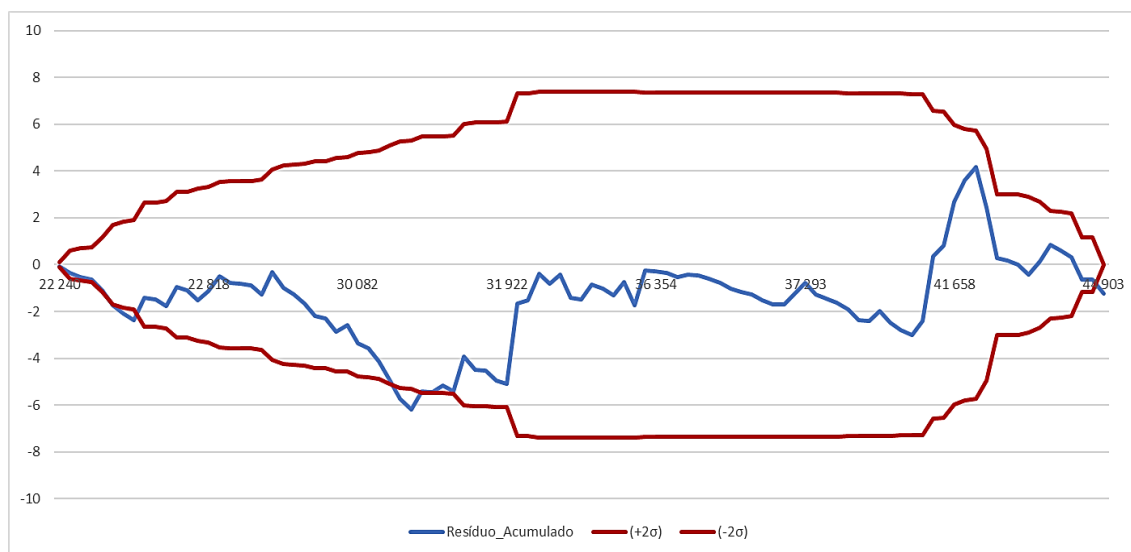


Figura 4.5: - Gráfico resíduos acumulados da variável TMDA para acidentes totais na A44 [2019]

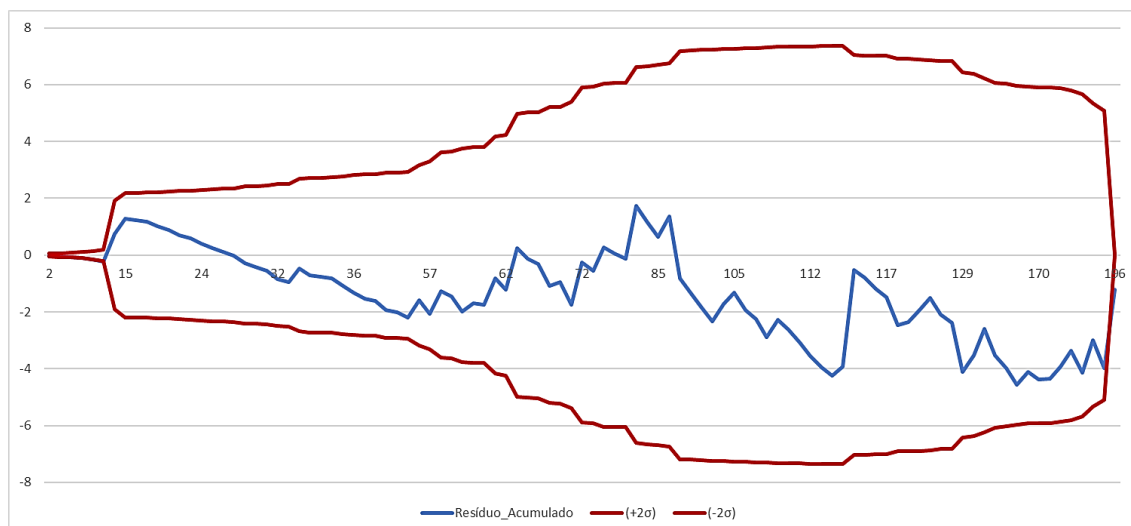


Figura 4.6: - Gráfico resíduos acumulados da variável comprimento para acidentes totais na A44 [2019]

Observando a figuras 4.5 é possível comprovar que os resíduos acumulados encontram-se distribuídos em torno do eixo horizontal, respeitando os limites definidos por $\pm 2\sigma$, exceto para valores de tráfego próximos de 31 000 e acima de 44000, onde ocorre uma extrapolação da banda delimitadora $\pm 2\sigma$, e não terminam no valor nulo. Deste modo, pode-se concluir que o modelo aplicado apresenta, em geral, um bom ajuste para o ano em análise, com exceção de alguns intervalos de valores do tráfego.

Para a variável comprimento (figura 4.6), os valores dos resíduos apresentam-se distribuídos em torno do eixo horizontal, respeitando os limites definidos por $\pm 2\sigma$, exceto para valores elevados de extensão de segmentos, que não termina no valor nulo. Assim, conclui-se que o modelo aplicado apresenta um bom ajuste, exceto para valores elevados de comprimento.

No que diz respeito aos gráficos de dispersão estes apresentam, por sua vez, uma relação com as variáveis contínuas de interesse - TMDA e COMPR. Podemos observar na figura 4.7 e na figura 4.8 a dispersão dos valores de acidentes observados e previstos, para as variáveis TMDA e o COMPR, respetivamente, para o ano de 2019.

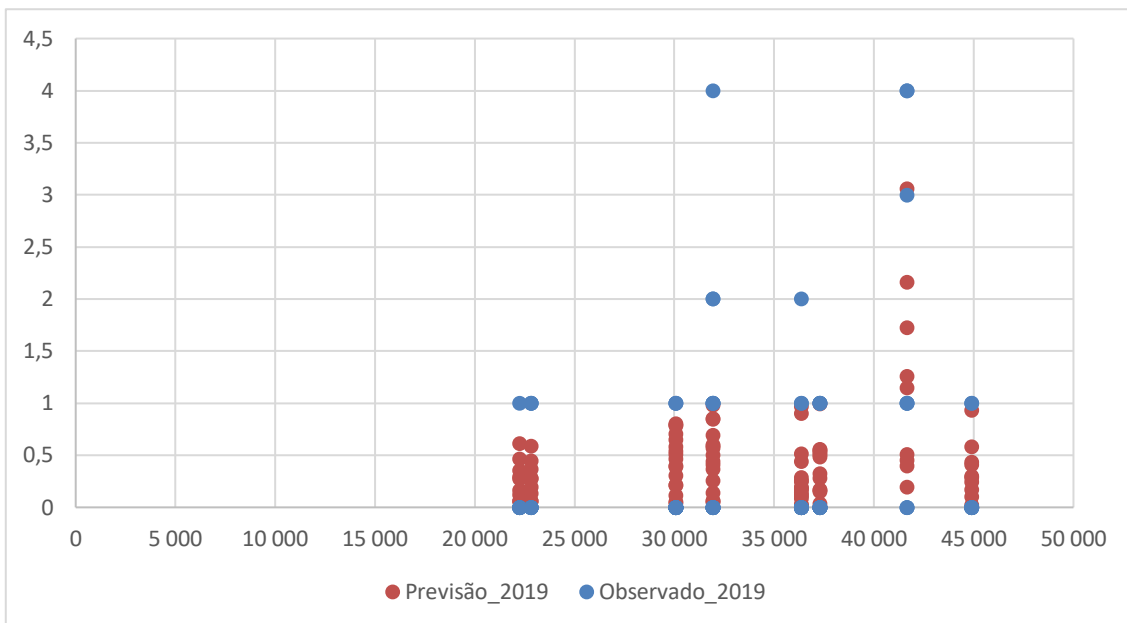


Figura 4.7: - Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A44 [2019]

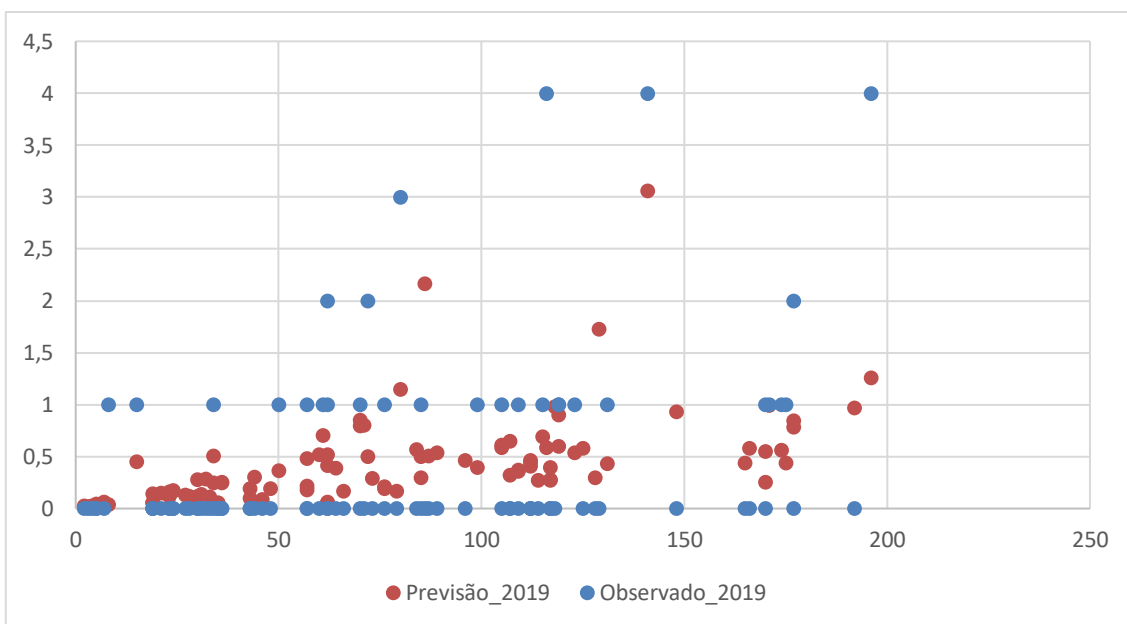


Figura 4.8: - Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A44 [2019]

Na figura 4.7 observa-se que os valores apresentam uma distribuição convergente, o que demonstra coerência para o modelo desenvolvido. No que diz respeito à variável comprimento, podemos observar, na figura 4.8, uma divergência de valores para extensões elevadas de segmentos. Analisando globalmente os valores, verifica-se uma distribuição convergente.

Relativamente ao ano de 2020, são apresentados os gráficos de resíduos acumulados para cada variável de exposição, nas figuras 4.9 e 4.10.

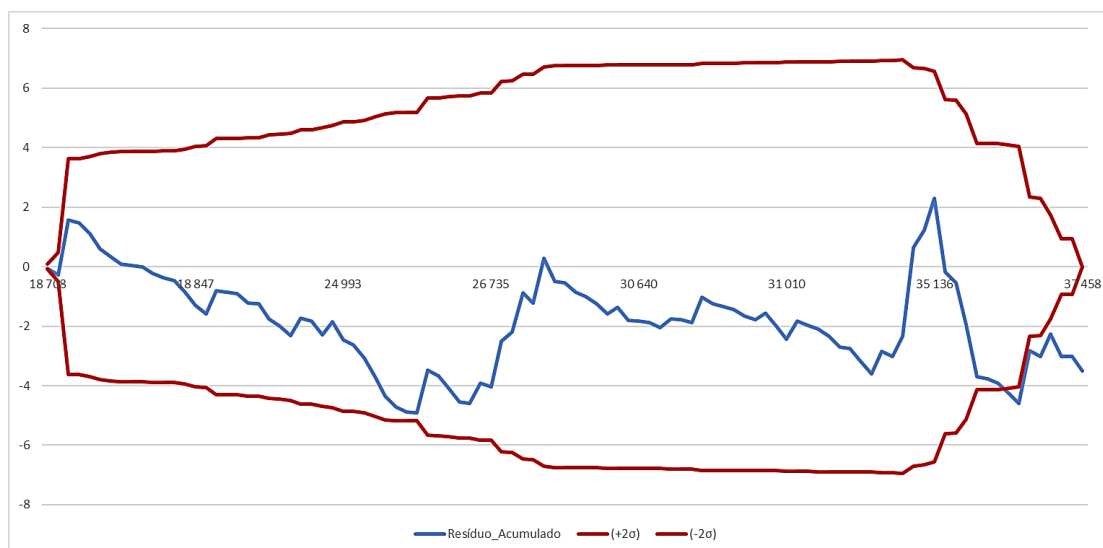


Figura 4.9: - Gráfico resíduos acumulados da variável TMDA para acidentes totais na A44 [2020]

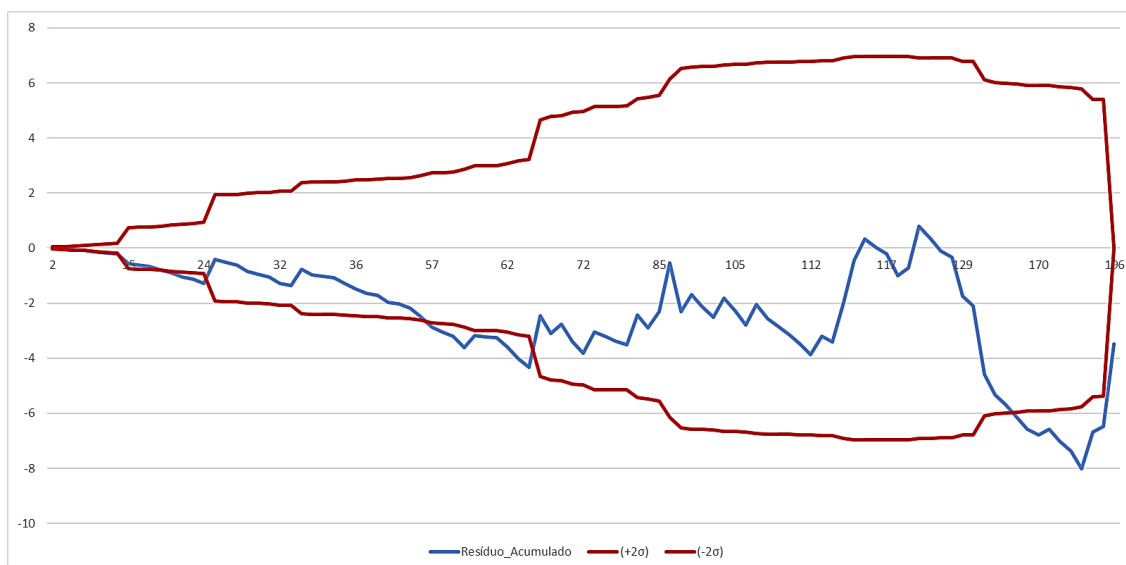


Figura 4.10: - Gráfico resíduos acumulados da variável comprimento para acidentes totais na A44 [2020]

O gráfico de resíduos acumulados em ordem ao TMDA da figura 4.9 mostra que a linha dos resíduos acumulados excede os limites definidos por $\pm 2\sigma$ para valores acima de 36000, não terminando no valor nulo. Deste modo, pode-se concluir que o modelo aplicado deixa de apresentar ajuste para valores elevados de tráfego.

Relativamente à variável comprimento (figura 4.10), é evidente uma oscilação dos resíduos acumulados. Para valores de extensão de segmento compreendidos entre 50 a 65m, assim como, extensões acima de 165m, observa-se uma extrapolação da banda $\pm 2\sigma$, não terminando no valor nulo. A partir destes factos, pode-se concluir que o modelo aplicado não se ajusta a todos os valores do comprimento, em especial para valores elevados.

Os gráficos de dispersão para o ano de 2020 estão presentes nas figuras 4.11 e 4.12 e permitem analisar os valores de acidentes observados e previstos, tendo em conta a variável de interesse - TMDA e COMPR, respetivamente.

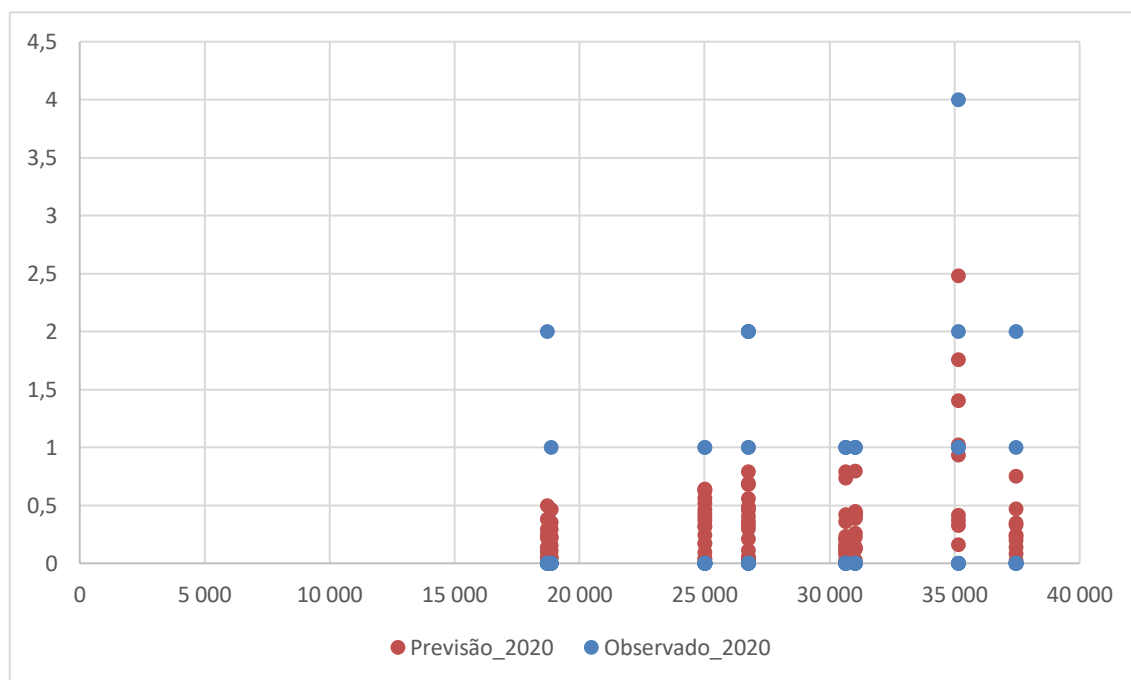


Figura 4.11: - Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A44 [2020]

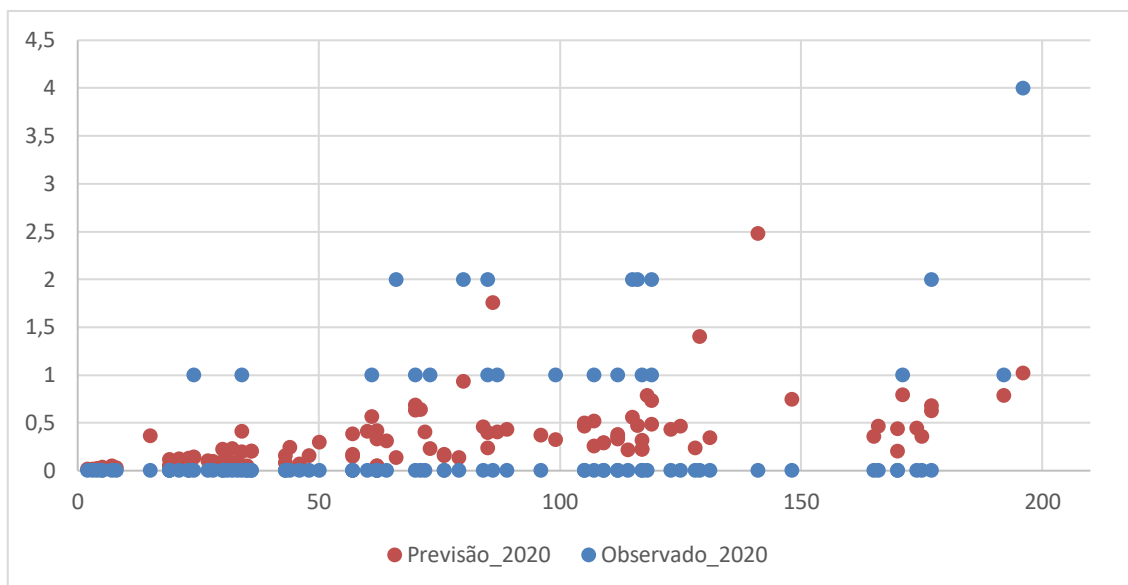


Figura 4.12: - Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A44 [2020]

Nas figuras 4.11 e 4.12 é visível a distribuição convergente dos valores previstos e observados, concluindo que o modelo demonstra coerência.

Modelo de Previsão de Acidentes com Vítimas – A44 (15-18)

Nas figuras seguintes são apresentados os gráficos de resíduos acumulados para cada variável de exposição, associado ao ano de 2019.

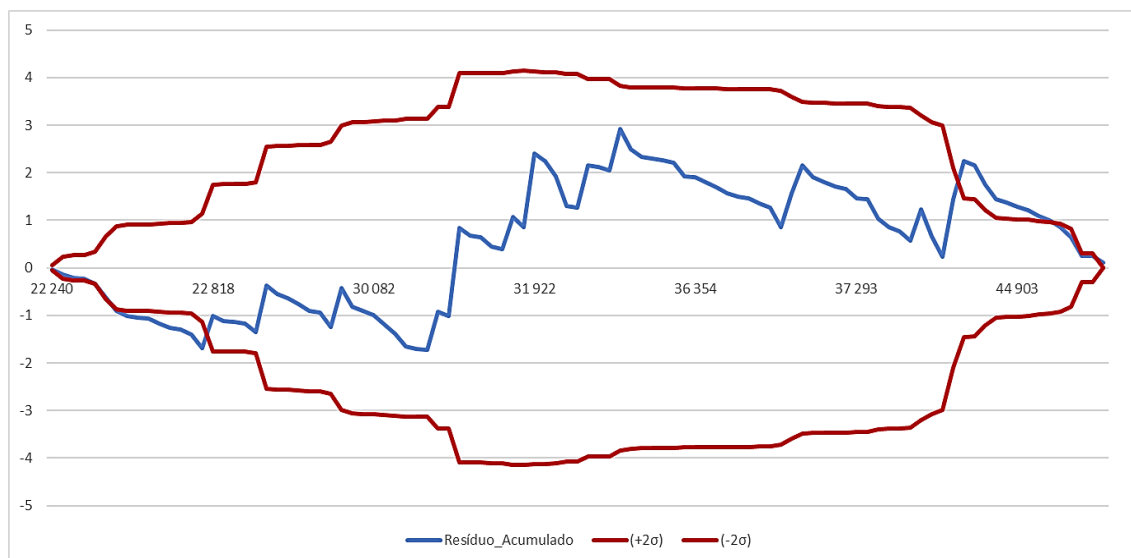


Figura 4.13: - Gráfico resíduos acumulados da variável TMDA para acidentes com vítimas na A44 [2019]

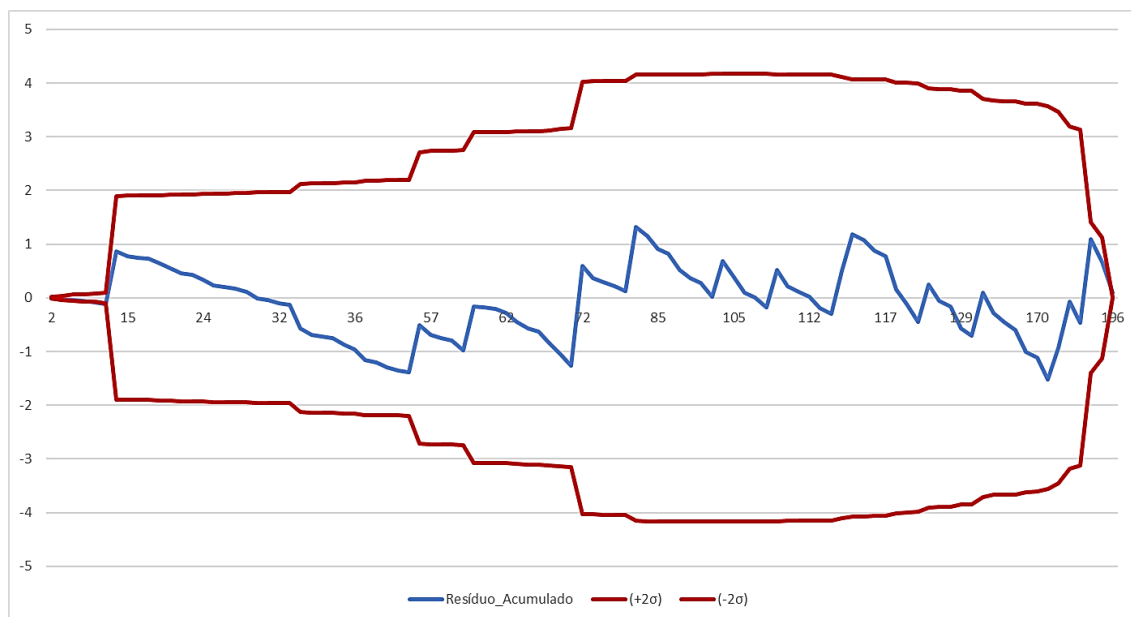


Figura 4.14: - Gráfico resíduos acumulados da variável comprimento para acidentes com vítimas na A44 [2019]

Na figura 4.13 podemos verificar que a linha dos resíduos acumulados excede os limites definidos por $\pm 2\sigma$ para os intervalos de tráfego entre 22300 e 22800 e acima de 41000 e não termina no valor nulo. Conclui-se que o modelo aplicado não se ajusta a todos os valores do tráfego, especialmente para valores elevados.

A variável comprimento, presente na figura 4.14, apresenta uma curva segundo o padrão desejado, ou seja, os resíduos acumulados encontram-se distribuídos em torno do eixo horizontal e estão contidos entre as bandas definidas por $\pm 2\sigma$. Deste modo, o modelo apresenta um bom ajuste para a variável comprimento.

Os gráficos de dispersão para o ano de 2019 estão presentes nas figuras 4.15 e 4.16 e permitem analisar os valores de acidentes observados e previstos, tendo em conta a variável de interesse - TMDA e COMPR, respetivamente.

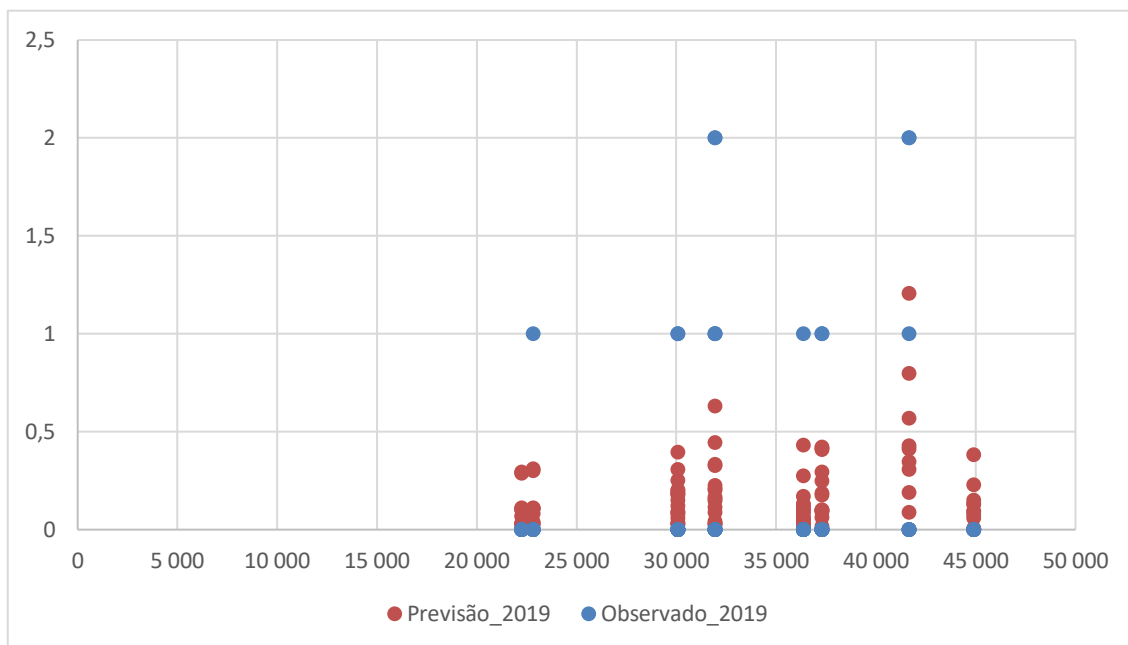


Figura 4.15: - Gráfico dispersão dos valores para a variável TMDA para acidentes com vítimas na A44 [2019]

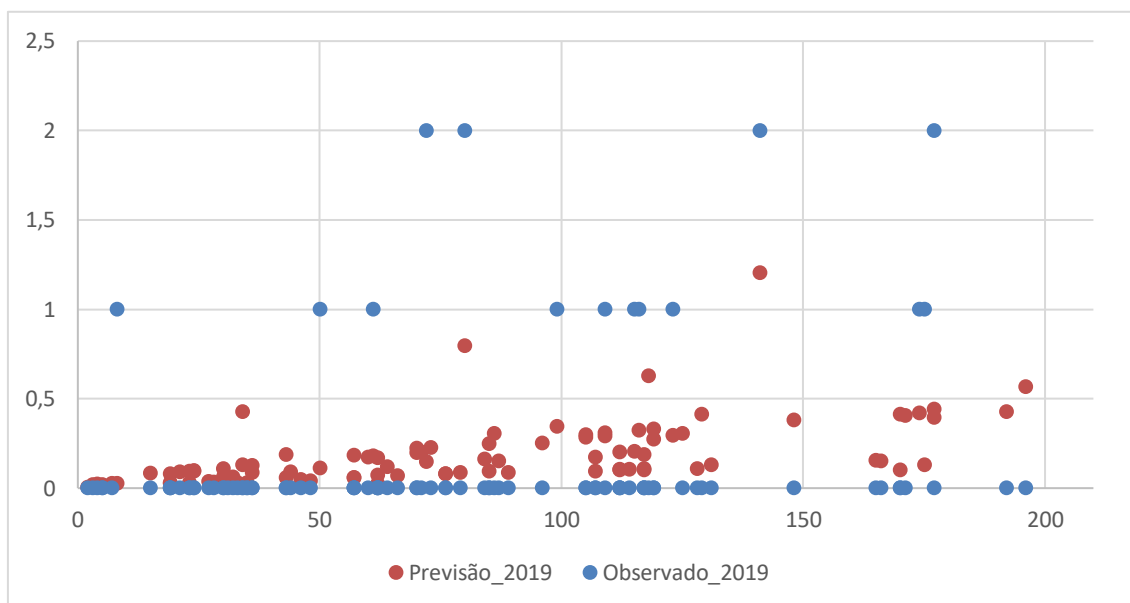


Figura 4.16: - Gráfico dispersão dos valores para a variável comprimento para acidentes com vítimas na A44 [2019]

Nas figuras 4.15 e 4.16, é visível a distribuição convergente dos valores previstos e observados, exceto para valores de extensão elevada, em que se nota uma divergência de valores. Analisando de uma forma geral, pode-se concluir que o modelo demonstra coerência.

No que diz respeito ao ano de 2020, os gráficos de resíduos acumulados, para cada variável de exposição, estão presentes nas figuras 4.17 e 4.18.

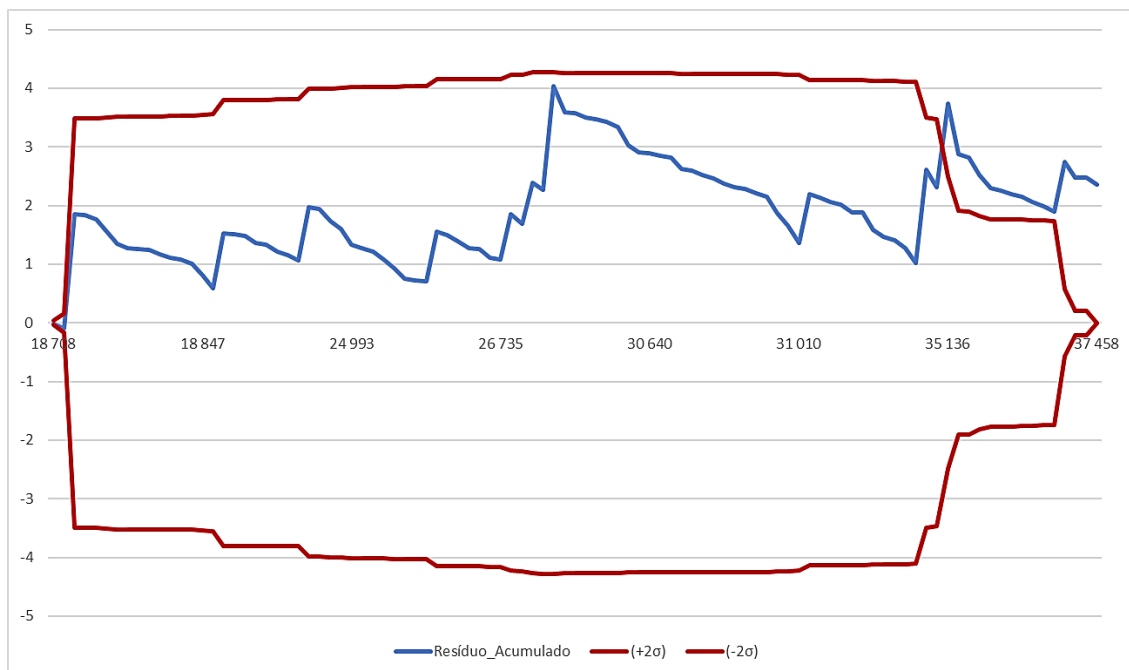


Figura 4.17: - Gráfico resíduos acumulados da variável TMDA para acidentes com vítimas na A44 [2020]

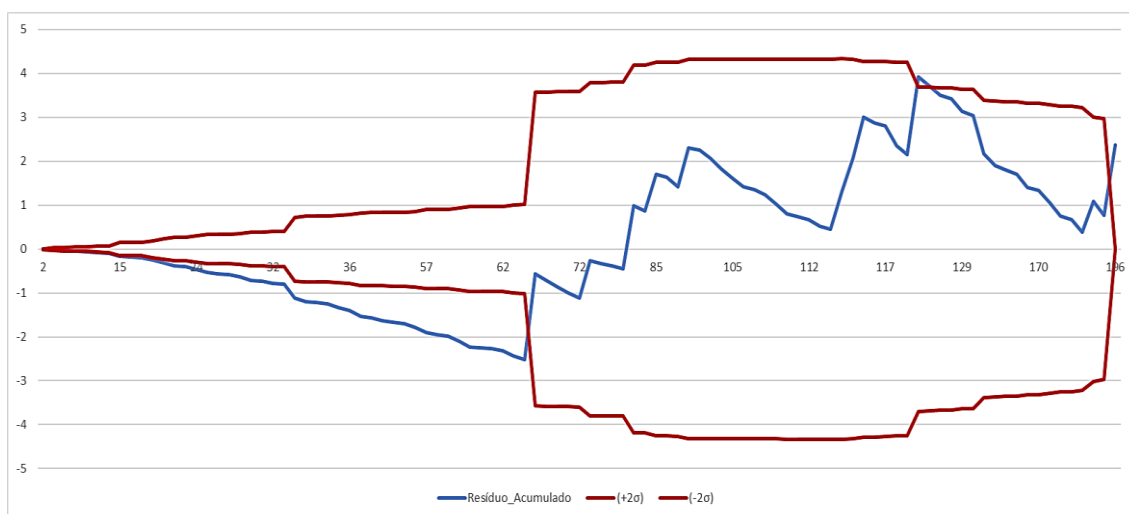


Figura 4.18: - Gráfico resíduos acumulados da variável comprimento para acidentes com vítimas na A44 [2020]

Na figura 4.17 pode-se verificar que a linha dos resíduos acumulados excede os limites definidos por $\pm 2\sigma$ para valores de tráfego acima de 35000 e não termina no valor nulo. O modelo aplicado demonstra um desajuste, em especial para valores elevados.

O gráfico de resíduos em ordem ao comprimento, presente na figura 4.18, mostra várias transposições do padrão desejado, ou seja, os resíduos acumulados não se encontram distribuídos em torno do eixo horizontal e ultrapassam as bandas definidas por $\pm 2\sigma$ nos intervalos de valores

de 10 a 65m, para valores próximos de 120m e acima de 190m. Deste modo, o modelo encontra-se desajustado para vários valores desta variável.

Os gráficos de dispersão para o ano de 2020 estão presentes nas figuras 4.19 e 4.20 e permitem analisar os valores de acidentes observados e previstos, tendo em conta a variável de interesse - TMDA e COMPR, respetivamente.

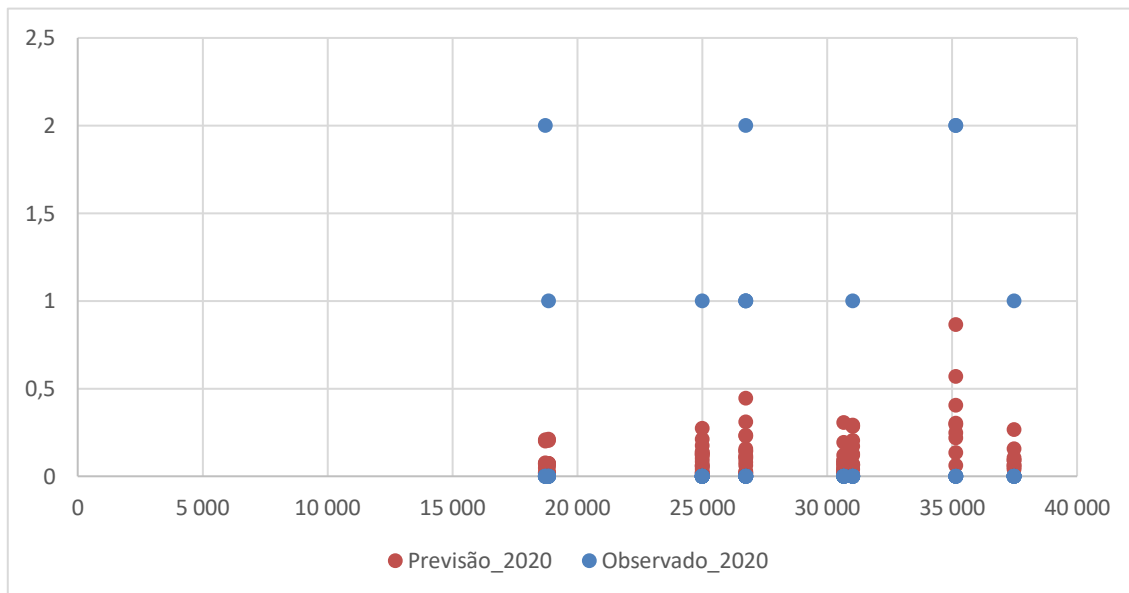


Figura 4.19: - Gráfico dispersão dos valores para a variável TMDA para acidentes com vítimas na A44 [2020]

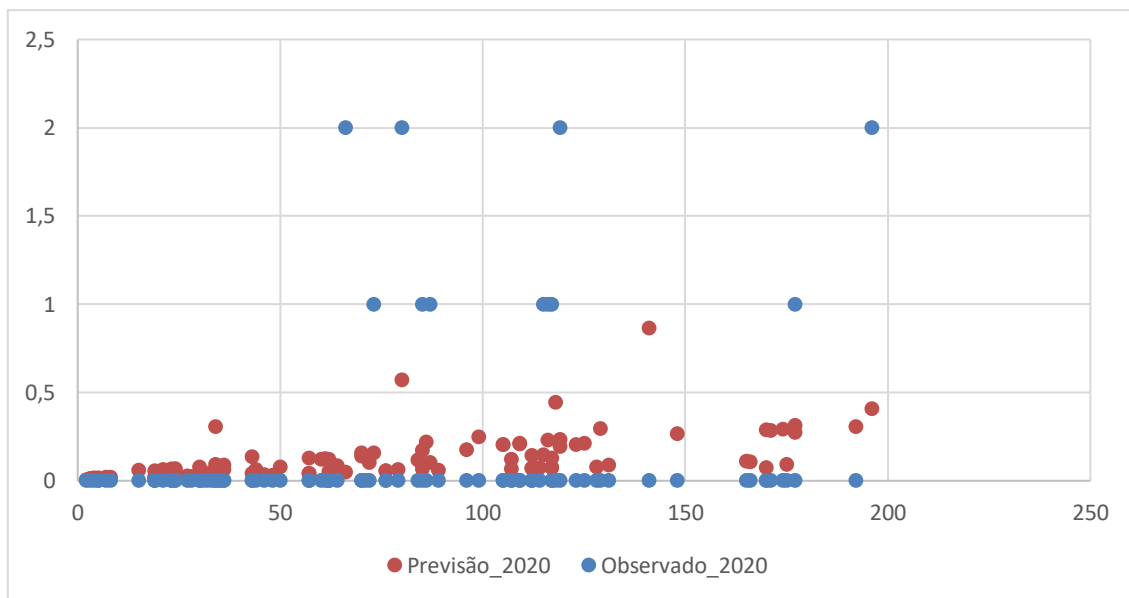


Figura 4.20: - Gráfico dispersão dos valores para a variável comprimento para acidentes com vítimas na A44 [2020]

Nas figuras 4.19 é visível uma distribuição convergente dos valores, concluindo que o modelo apresenta coerência. Já na figura 4.20 pode-se observar uma distribuição divergente dos valores previstos para extensões de segmentos próximas de 140m, em relação aos valores observados, deduzindo-se que o modelo demonstra falta de coerência para estes valores.

Modelo de Previsão de Acidentes Totais- A29 (15-17)

Nas figuras seguintes são apresentados os gráficos de resíduos acumulados para cada variável de exposição, associado ao ano de 2019.

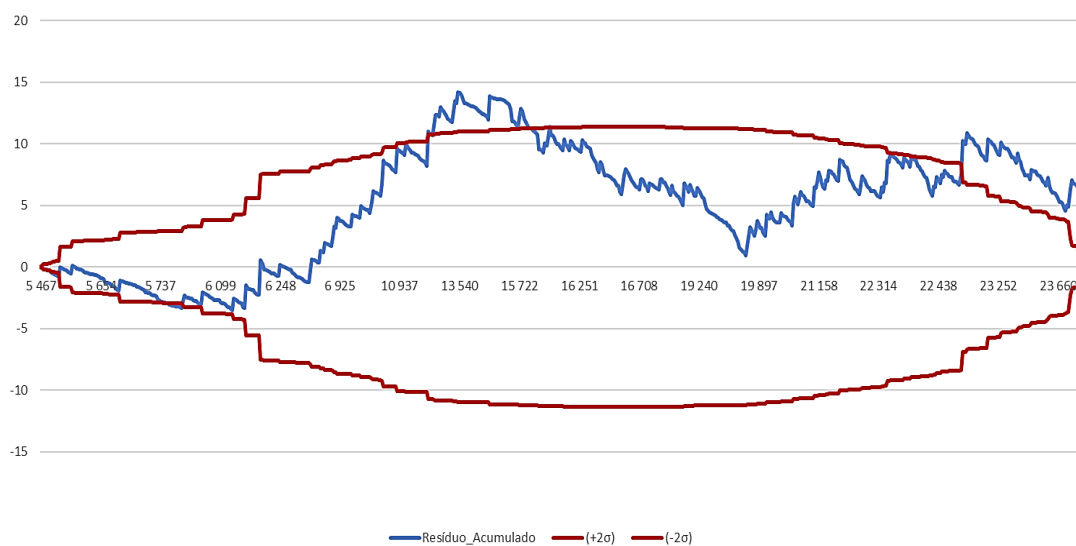


Figura 4.21: - Gráfico resíduos acumulados da variável TMDA para acidentes totais na A29 [2019]

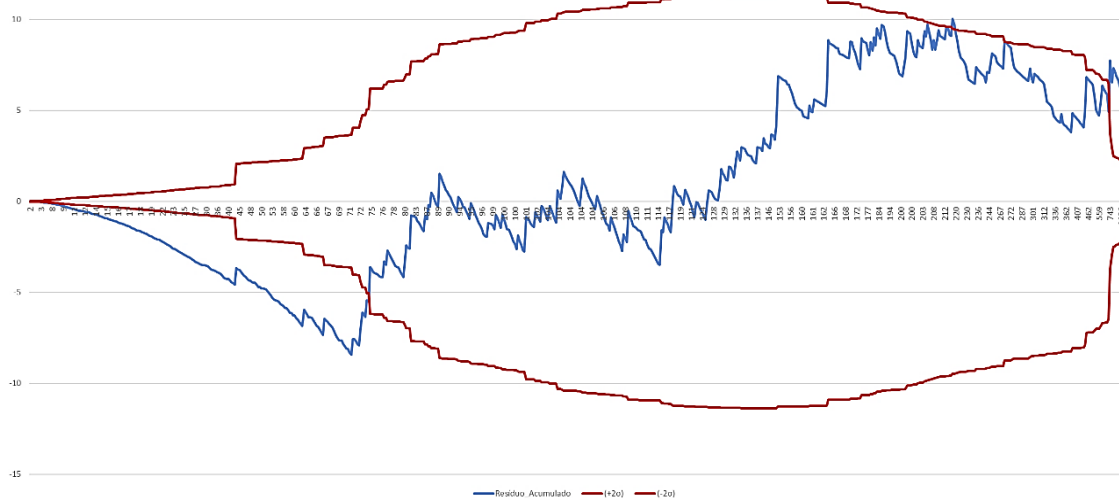


Figura 4.22: - Gráfico resíduos acumulados da variável comprimento para acidentes totais na A29 [2019]

O gráfico da figura 4.21 relativo à variável TMDA, mostra que a linha dos resíduos acumulados excede os limites definidos por $\pm 2\sigma$ para o intervalo de valores de tráfego entre 11600 e 15800 e acima de 22500, não terminando no valor nulo. O modelo aplicado demonstra um desajusta, em especial para valores elevados.

Na figura 4.22, está presente o gráfico relativo à variável comprimento onde é evidente a extrapolação da linha dos resíduos acumulados em relação à banda $\pm 2\sigma$, para os intervalos de valores de extensão entre 8m e 75m e acima de 650m. Deste modo, o modelo encontra-se desajustado para vários valores desta variável, especialmente para valores elevados de comprimento.

Nas figuras 4.23 e 4.24 estão presentes os gráficos de dispersão para o ano de 2019, que permitem analisar os valores de acidentes observados e previstos, tendo em conta a variável de interesse - TMDA e COMPR, respetivamente.

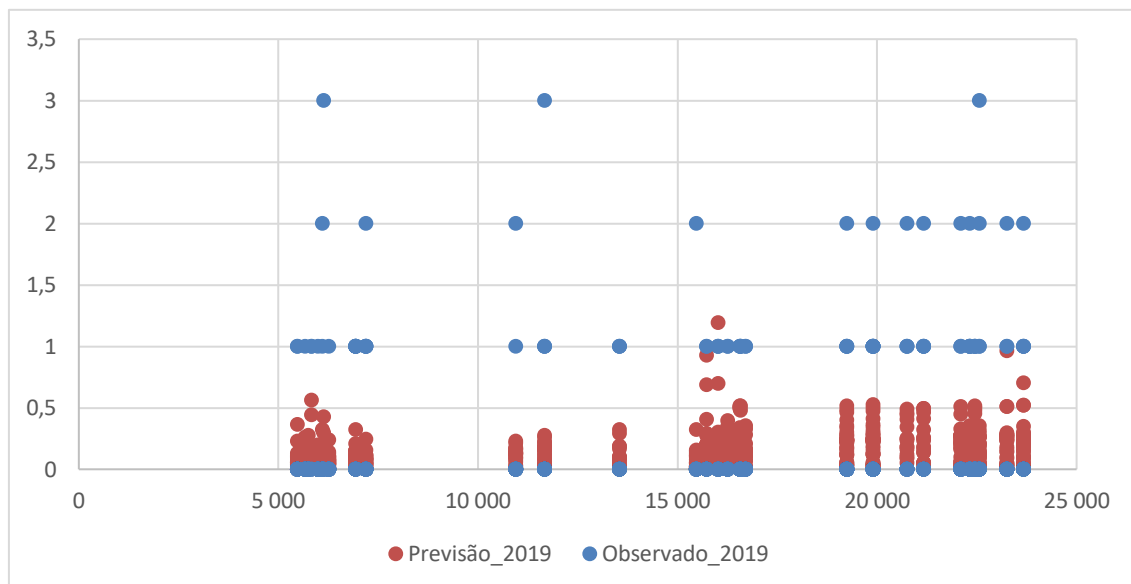


Figura 4.23: - Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A29 [2019]

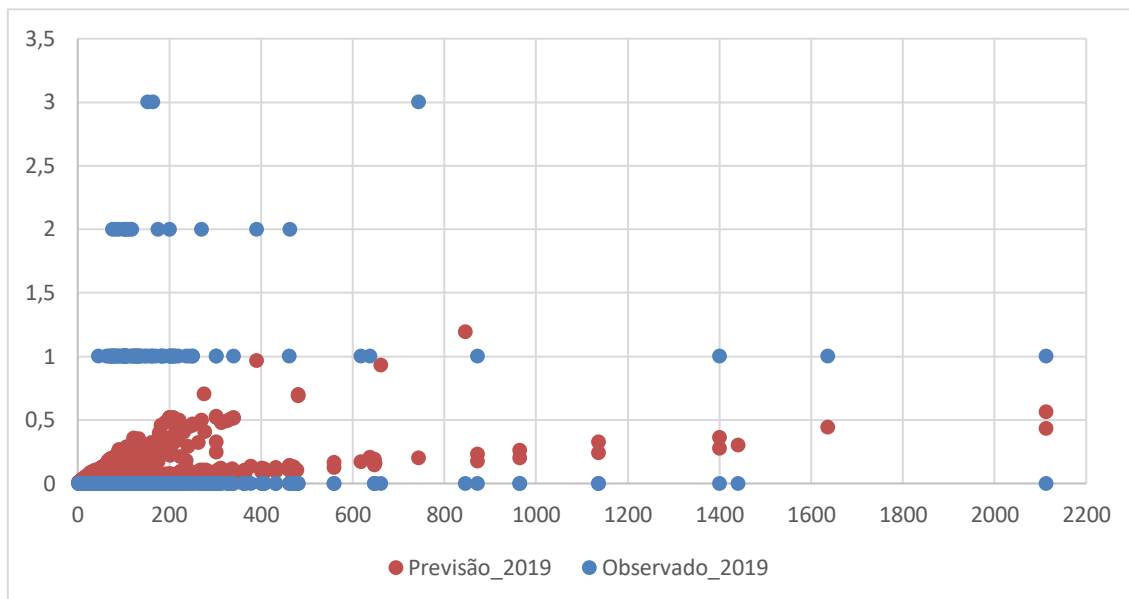


Figura 4.24: - Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A29 [2019]

Nos gráficos das figuras 4.23 e 4.24, é visível a distribuição convergente dos valores, concluindo que o modelo demonstra coerência.

Os gráficos de resíduos acumulados para cada variável resposta, referentes ao ano de 2020, estão apresentados, nas figuras 4.25 e 4.26.

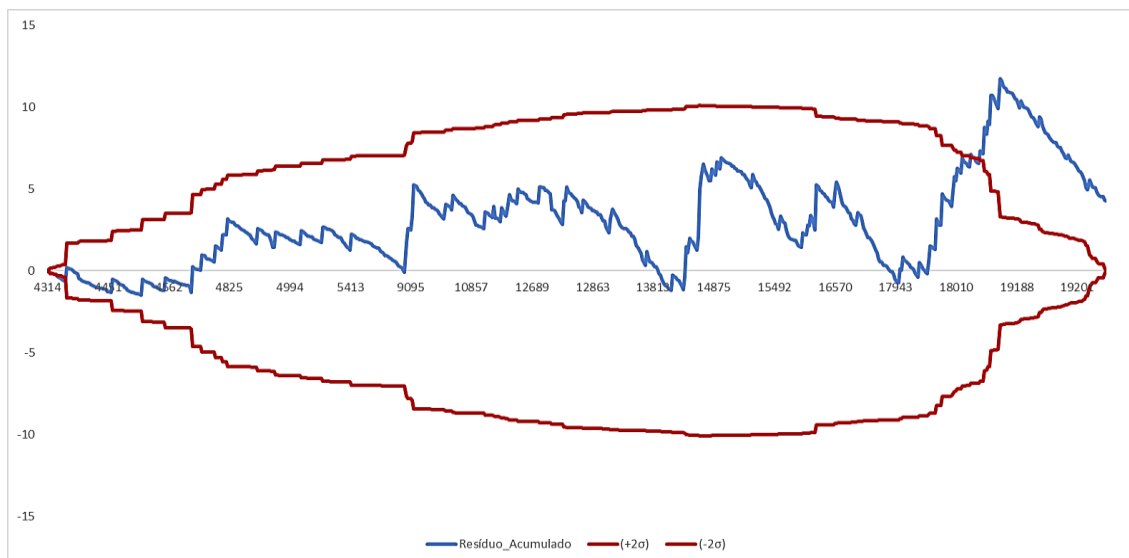


Figura 4.25: - Gráfico resíduos acumulados da variável TMDA para acidentes totais na A29 [2020]

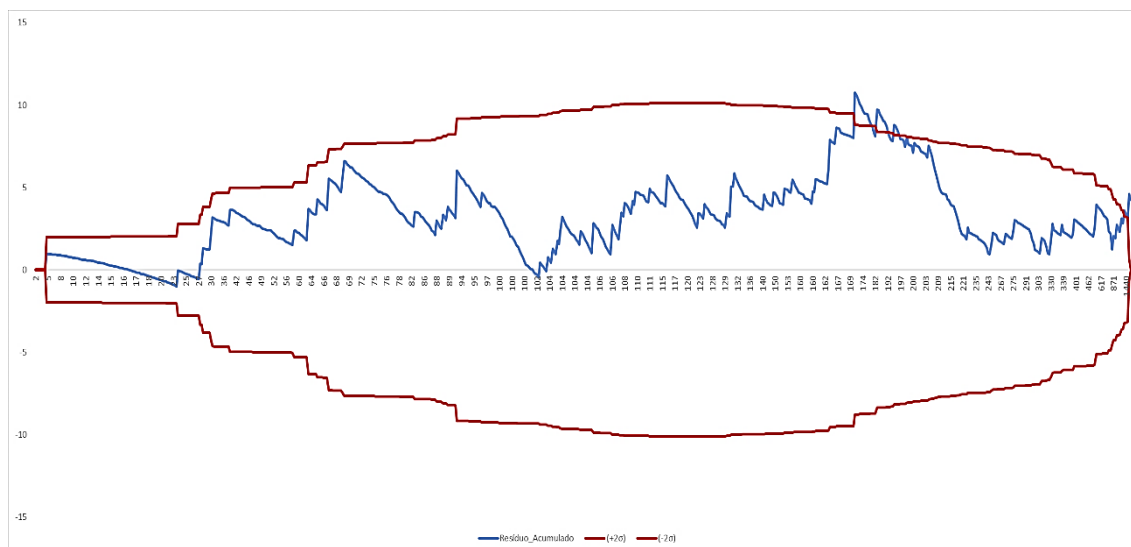


Figura 4.26: - Gráfico resíduos acumulados da variável comprimento para acidentes totais na A29 [2020]

Na figura 4.25 está presente o gráfico relativo à variável TMDA, demonstrando que a curva de resíduos acumulados ultrapassa os limites definidos por $\pm 2\sigma$, para valores de tráfego elevados, ou seja, acima de 18800, não terminando no valor zero. Como tal, o modelo aplicado não se ajusta para valores elevados de tráfego. O gráfico da variável comprimento, presente na figura 4.26, apresenta uma curva segundo o padrão desejado, exceto para valores de extensão compreendidos entre 170m e 200m, e acima de 1400m, em que a linha de resíduos acumulados excede os limites das bandas definidas por $\pm 2\sigma$, não coincidindo com o valor nulo final. Deste modo, o modelo apresenta um desajuste para alguns valores de extensão.

Nas figuras 4.27 e 4.28 estão presentes os gráficos de dispersão para o ano de 2020 e permitem analisar os valores de acidentes observados e previstos, tendo em conta a variável de interesse - TMDA e COMPR, respetivamente.

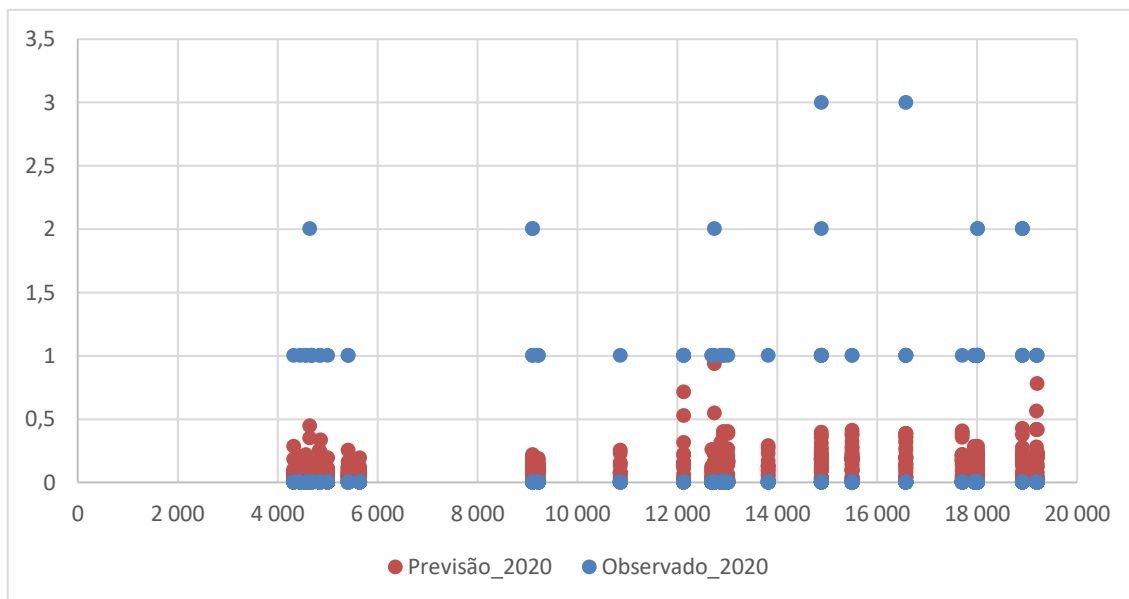


Figura 4.27: - Gráfico dispersão dos valores para a variável TMDA para acidentes totais na A29 [2020]

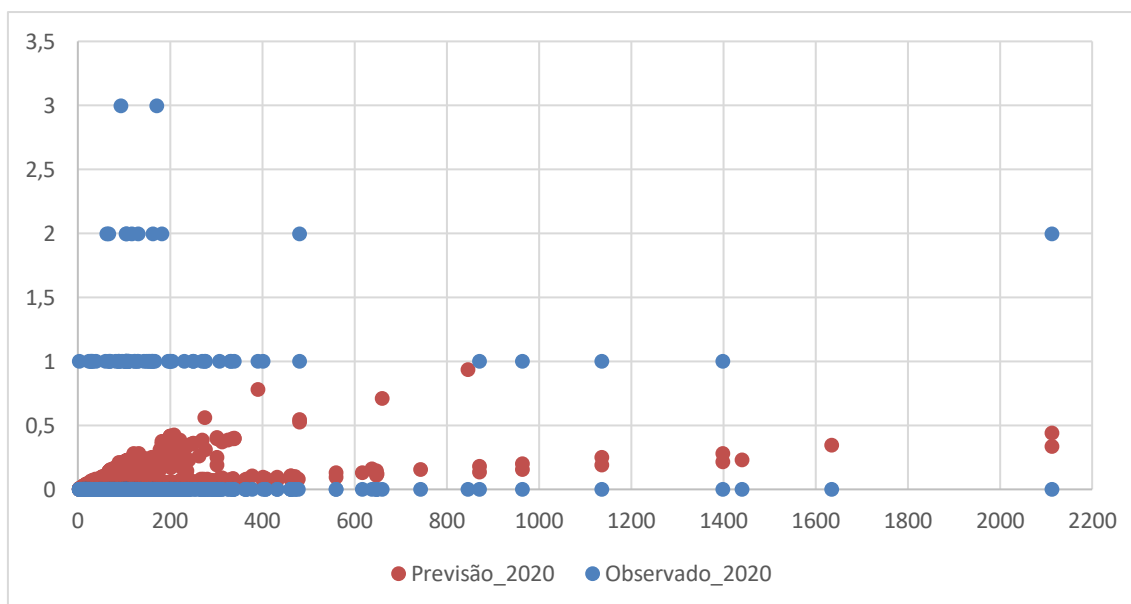


Figura 4.28: - Gráfico dispersão dos valores para a variável comprimento para acidentes totais na A29 [2020]

Pelos gráficos de dispersão, presentes nas figuras 4.27 e 4.28, é possível verificar uma distribuição convergente dos valores, concluindo que o modelo demonstra coerência.

5

CONCLUSÕES

A sinistralidade rodoviária é um problema mundial, que deve ser enfrentado através estratégias que visem um sistema de mobilidade segura, dada a inevitabilidade do erro humano.

Os modelos de previsão de acidentes têm-se demonstrado vantajosos na análise das causas dos acidentes, principalmente, no que toca à infraestrutura. Estes permitem estabelecer a relação entre os fatores de risco com o número de acidentes ocorridos num determinado espaço e tempo. Assim, é possível aplicar medidas corretivas, mesmo antes de uma ação acontecer, tendo por base estudos fundamentados.

Os MPAs mais comumente aplicados seguem uma distribuição de Poisson ou BN. Neste trabalho, a componente aleatória segue uma distribuição de Poisson, tanto para a previsão de frequência de acidentes totais, como para acidentes com vítimas.

Ao longo do trabalho, foi realizado um estudo com o objetivo de desenvolver modelos de previsão de frequência de acidentes para a autoestrada A44, da concessionária Ascendi. O processo teve como etapas, a identificação dos dados relativos a acidentes e características geométrica, operacionais e do pavimento, obtidas através das bases de dados da empresa, que foram posteriormente convertidas em variáveis úteis para o desenvolvimento dos modelos.

Os MPAs foram desenvolvidos para o período de anos de 2015 a 2020 e alcançaram resultados satisfatórios, que permitem corroborar a elevada influência das variáveis de exposição, TMDA e comprimento, no número de acidentes conforme referenciado nos estudos.

Ainda, foi possível estudar o efeito que as restantes variáveis de exposição produziam, no número de acidentes, para cada modelo. Assim, são descritos os fatores que apresentam um efeito preponderante, quer positivo, quer negativo, na ocorrência de acidentes para os modelos:

Modelo de acidentes totais:

- Limite de velocidade de 80km/h - representa um aumento de 47% nos acidentes, em relação ao limite de velocidade de 100km/h;
- Raio de curva inferior a 1000m – representa uma diminuição 253% nos acidentes, em relação a segmento em reta.

Modelo de acidentes com vítimas:

- TMDA – representa um aumento de 22% no número de acidentes;
- Raio de curva inferior a 1000m – representa uma diminuição 592% nos acidentes, em relação a segmento em reta.

Na etapa de avaliação, procedeu-se à análise de resíduos com o objetivo de aferir a qualidade de ajuste do modelo aos dados. Foram elaborados gráficos de resíduos acumulados dos acidentes ordenados de acordo com o TMDA e o comprimento, concluindo-se que, de uma forma geral, o modelo apresenta um bom desempenho para ambos os modelos da A44.

Para melhor aferir a aplicação de MPA, foram estudados três modelos distintos, para duas autoestradas distintas, considerando apenas a previsão para os anos 2019 e 2020, de forma a serem comparáveis os resultados das duas autoestradas. Assim, foram aplicadas diversas medidas de desempenho, assim como, análises gráficas de qualidade de ajuste. Esta análise comparativa permitiu concluir que, apesar das alterações verificadas na mobilidade para o ano 2020, os modelos conseguiram estabelecer a mesma fiabilidade, em relação a um ano comum em termos de mobilidade (2019). Os gráficos de resíduos acumulados para os anos referidos apresentam-se ajustados, com exceção para valores elevados das variáveis de exposição, TMDA e comprimento, que os modelos revelam desajuste.

Limitações

Uma limitação do estudo foi a organização das bases de dados existentes, que não se apresentam organizadas para a finalidade deste estudo. Após ultrapassar este problema, com o apoio do Departamento de Informática da Ascendi, detetou-se a omissão de dados nas bases de dados existentes. Ainda, não foi possível analisar outras variáveis com relevância para o estudo como,

por exemplo, a velocidade operacional. No caso da velocidade operacional, é necessário uma recolha extensiva nos segmentos em análise.

Recomendações para trabalhos futuros

Recomenda-se a correção dos erros detetados nas bases de dados da Ascendi, principalmente a omissão de dados em alguns sublanços, que impede a segmentação da rede, importante para o desenvolvimento dos modelos, assim como, melhorar a uniformização dos termos utilizados nas bases de dados em bruto.

Ainda, sugere-se a inclusão de outras variáveis independentes, que caracterizam de forma operacional as infraestruturas, como a iluminação, as condições do pavimento, a sinalização, a velocidade operacional, entre outras. Por sua vez, é importante realizar o estudo tendo em conta outras variáveis dependentes como as vítimas e os diferentes tipos de lesão associados.

Finalmente, recomenda-se a atualização dos modelos existentes, através da introdução dos dados dos acidentes, assim como do volume de tráfego, anualmente, para futura aplicação. Este processo é finalizado após a introdução da base de dados atualizada, no programa NLOGIT, que produz um novo modelo mais robusto e eficaz.

REFERÊNCIAS BIBLIOGRÁFICAS

- AASHTO (2010) American Association of State and Highway Transportation Officials. *Highway Safety Manual*. Washington, USA.
- ANSR (2020). *Contributos para a visão zero 2030 – até 30 de Outubro*. <http://www.ansr.pt/Noticias/Pages/Estratégia-Nacional-de-Segurança-Rodoviária-2021-2030.aspx>.
- ARMSTRONG, J. S. (1985) Long-range Forecasting: From Crystal Ball to Computer, 2nd. ed. Wiley. ISBN 978-0-471-82260-8
- BENNETT, N. D., CROKE, B. F. W., GUARISO, G., GUILLAUME, J. H. A., HAMILTON, S. H., JAKEMAN, A. J., MARSILI-LIBELLI, S., NEWHAM, L. T. H., NORTON, J. P., PERRIN, C., PIERCE, S. A., ROBSON, B., SEPPELT, R., VOINOV, A. A., FATH, B. D. & ANDREASSIAN, V., (2013). *Characterising performance of environmental models*. Environmental Modelling & Software, v. 40, p. 1-20.
- BEN HAMNER, MICHAEL FRASCO & ERIN LEDELL (2018) *Evaluation Metrics for Machine Learning*. <https://github.com/mfrasco/Metrics>
- BIONDI, D., FRENI, G., IACOBELLIS, V., MASCARO, G. & MONTANARI, A. (2012) *Validation of hydrological models: conceptual basis, methodological approaches and a proposal for a code of practice*. Physics and Chemistry of the Earth, Parts A/B/C, v. 4244, p. 70-76.
- BONNESON, J.A., PRATT, M.P., (2008). *Procedure for developing accident modification factors from cross-sectional data*. Transportation Research Record 2083, 40-48.
- BORSOS, A., JOHN, N. I., GYULA, O. (2014) *Development of safety performance functions for two-lane rural first-class main roads in Hungary*. Proceedings of 5th Transport Research Arena Conference, Paris, France.
- CAFISO, S., DI GRAZIANO, A., DI SILVESTRO, G., LA CAVA, G. (2008) *Safety Performance Indicators for Local Rural Roads: Comprehensive Procedure from Low-Cost Data Survey to Accident Prediction Model*. Proceedings of the 87th Annual Meeting of the Transportation Research Board, Washington, USA.
- CARDOSO, J. L. (1998). *Definição e deteção de zonas de acumulação de acidentes*. In: Segurança Rodoviária - Avaliação e Redução da Sinistralidade, Lisboa.
- CARDOSO, J. (2007). *Métodos racionais de apoio à intervenção da engenharia em segurança rodoviária*. Programa de investigação para obtenção do título "Habilitado para o exercício de funções de Coordenação de Investigação Científica", LNEC, Lisboa.

- CARDOSO, J. (2009). *Modelos de estimativa de frequência de acidentes e vítimas para dados bidirecionais de autoestradas portuguesas – uma abordagem bayesiana*. Lisboa: LNEC, 5/2010 – NPTS.
- COMISSÃO EUROPEIA, (2021). *Road safety: 4,000 fewer people lost their lives on EU roads in 2020 as death rate falls to all-time low*. Brussels, 20 Abril.
- CHANG, L. (2005). *Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network*. Safety Science, v. 43, p. 541-557.
- CHANG, L. & WANG, H. (2006). *Analysis of traffic injury severity: An application of non-parametric classification tree techniques*. Accident Analysis and Prevention, v. 38, p. 1019-1027.
- CLAUDE, Giovana (2012) – *Previsão da ocorrência de acidentes de trânsito em interseções de vias arteriais urbanas – o caso de Taguatinga /DF*. Departamento de Engenharia Civil e Ambiental, Universidade de Brasília. Dissertação de Mestrado em Transportes.
- COOK, D., REGINALD, S. & JUSTIN, J. (2011). *Effect of Road Segmentation on Highway Safety Analysis*. Proceedings of the 90th Annual Meeting of the Transportation Research Board, Washington, USA.
- CRONBACH, L. J. (1971) *Test Validation*. Educational Measurement. 2nd ed., R. L. Thorndike (ed.), American Council on Education, Washington, USA, p. 443–507.
- CUNTO, F. J. C. (2009). Projeto de Pesquisa: Modelos de Previsão de Acidentes de Trânsito em Vias urbanas brasileiras – Edital MCT/CNPq 18/2009 – UFC - Fortaleza - CE.
- DAS, A. & ABDEL-ATY, M. (2010) *A genetic programming approach to explore the crash severity on multi-lane roads*. Accident Analysis and Prevention, v. 42, p. 548-557.
- DAS, A. & ABDEL-ATY, M. (2011). *A combined frequency-severity approach for the analysis of rear-end crashes on urban arterials*. Safety Science, v. 49, p. 1156-1163.
- ELVIK, R. (2009) *The power model of the relationship between speed and road safety*. Update and new analyses. Report 1034/2009, Institute of Transport Economics, Oslo, Denmark.
- FLORES, B. E. (1986) "Uma visão pragmática de medição de precisão em previsão", Omega (Oxford), 14 (2), 93-98. doi:10.1016/0305-0483(86)90013-7
- FRIDSTROM, L., et al. (1995). *Measuring the contribution of randomness, exposure, weather and daylight to the variation in road accident counts*. Accident Analysis and Prevention, 27, 1-20.

- GLENNON, J. C. (1987) *Effect of Alignment on Highway Safety*. State of the Art Report 6 – Relationships between Safety and Key Highways Features. Proceedings of the 66th Annual Meeting of the Transportation Research Board, Washington, USA.
- GOLOB, T., RECKER, W. and ALVAREZ, V. (2004). *Freeway safety as a function of traffic flow*. Accident Analysis and Prevention, 36, 933-946.
- GOMES, S. C. G. V. (2010). *Avaliação da Influência da Infraestrutura na Segurança Rodoviária em Meio Urbano*. Tese de Doutorado, Publicação, Departamento de Engenharia Civil, Universidade de Coimbra, Coimbra – Portugal, 404 Pg.
- GREEN, E. R. & AGENT, K. R. (2011) *Evaluation of the Locations of Kentucky's Traffic Crash Data*. Proceedings of the 3rd International Conference on Road Safety and Simulation, Indianapolis, USA.
- GREEN, E., NIKIFOROS, S. & REGINALD, S. (2017). *Segment Length and Highway Safety Analysis: Does It Matter?*. Proceedings of the 96th Annual Meeting of the Transportation Research Board, Washington, USA.
- GREENE, W. H. (2008). *Econometric Analysis*, Pearson International Edition, New Jersey.
- GREIBE, P. (2003). *Accident prediction models for urban roads*. Accident Analysis and Prevention, 35, 273-285.
- GUIMARÃES, R. C. e CABRAL, J. S. (1997). *Estatística*, McGraw-Hill, Lisboa etc.
- HAUER, E. (2004) *Statistical Road Safety Modeling*. Transportation Research Record, n. 1897, p. 81-87.
- HAUER, E. e BAMFO, J. (1997). *Two tools for finding what function links the dependente variable to the explanatory variables*. In: ICTCT, Lund.
- HUMPHREY, G. B., MAIER, H. R., WU, W., MOUNT, N. J., DANDY, G. C., ABRAHART, R. J. & DAWSON, C. W. (2017) *Improved validation framework and R-package for artificial neural network models*. Environmental Modelling & Software, v. 92, p. 82-106.
- ITF (2008) International Transport Forum. *Towards Zero: Ambitious Road Safety Targets and the Safe System Approach*. Research Report, OECD Publishing, Paris, France.
- KIM, D. & WASHINGTON, S. (2006) *The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models*. Accident Analysis and Prevention, v. 38, p. 1094-1100.

- KOCKELMAN, K.M. e KWEON, Y. J. (2002). *Driver injury severity: an application of ordered probit models*. Accident Analysis and Prevention, 34, 313-321.
- LORD, D. (2000) *The prediction of accidents on digital networks: characteristics and issues related to the application of accident prediction models*. PhD, Department of Civil Engineering, University of Toronto, Toronto
- LORD, D. (2006). *Modeling motor vehicle crashes using Poisson-gamma models: examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter*. Accident Analysis and Prevention, 38, 751-766.
- LORD, D. e PARK, P. Y. J. (2008). *Investigating the effects of the fixed and varying dispersion parameters of Poisson - gamma models on empirical Bayes estimates*. Accident Analysis and Prevention, 40, 1441-1457.
- LORD, Dominique e MANNERING, Fred (2010) - *The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives*. Transportation research part A: policy and practice.
- LU, J., ALBERT, G., KIROLOS, H. & WANYANG, W. (2013). *Clustering-based roadway segment division for the identification of high-crash locations*. Journal of Transportation Safety & Security, v. 5(3), p. 224-239.
- MADALOZO, H. C. (2003) *Desenvolvimento de ferramenta de análise de curvas horizontais de rodovias, para melhoria de projeto e operação, utilizando redes neurais artificiais*. Dissertação de Mestrado em Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná, Curitiba.
- MAHER, M. e SUMMERSGILL, I. (1995). *A comprehensive methodology for the fitting of predictive accident models*. Accident Analysis and Prevention, 28, 281-296.
- MAHER, M. J.; SUMMERSGILL, I. (1996). *A Comprehensive Methodology for the Fiting of Predictive Accident Models – Accident Analysis and Prebention*, Reino Unido, v. 28, n. 3 p. 281-296.
- MIAOU, S.P., LORD, D., (2003). *Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus Empirical Bayes*. Transportation Research Record 1840, 31-40.
- MAYCOCK, G. e HALL, R. D. (1984) *Accidents at 4-Arm Roundabouts*. TRL1120, Transport and Road Research Laboratory (TRRL), Crowthorne, Inglaterra.

- MCCULLAGH, P.; NELDER, J. A. (1989). *Generalized Linear Models*, Chapman & Hall, London, etc.
- MOUNTAIN, L., FAWAZ, B. e JARRETT, D. (1996). *Accident prediction models for roads with minor junctions*. Accident Analysis and Prevention, 28, 695-707.
- ODGEN, K. W. (1996) *Safer roads: a guide to road safety engineering*. Ashgate Publishing limited, University Press, Cambridge, UK, 516 p.
- OCDE (1997). *Road safety principles and models: review of descriptive, predictive, risk and accident consequence models*. Paris: Organisation for Economic Co-operation and Development, OCDE/GD (97)153.
- OH, J., LYON, C., WASHINGTON, S., PERSAUD, B. & BARED, J. (2003) *Validation of FHWA Crash Models for Rural Intersections: Lessons Learned*. Transportation Research Record, n. 1840, p. 41-49.
- PEDHAZUR, E. J. & SCHMELKIN, L. P. (1991) *Measurement, Design, and Analysis: An Integrated Approach*. Lawrence Erlbaum Associates, Hillsdale, USA, 819 p.
- POLUS, A., POLLATSCHEK, M. and FARAH, H. (2005). *Impact of infrastructure characteristics on road crashes on two-lane highways*. Traffic Injury Prevention, 6, 240-247.
- POWER, M. (1993) *The predictive validation of ecological and environmental models*. Ecological Modelling, v. 68, p. 33-50.
- QIN, X. & WELLNER, A. (2012) *Segment Length Impact on Highway Safety Screening Analysis*. Proceedings of the 91st Annual Meeting of the Transportation Research Board, Washington, USA.
- RESENDE, P. & BENEKOHAL, R. (1997) *Effect of roadway section length on accident modelling traffic congestion and traffic safety*. Proceedings of 21st Century Conference, Chicago, USA.
- REURINGS, M., et al. (2005). "Accident Prediction Models and Road Safety Impact Assessment: a state-of-the-art." RIPCORD-ISEREST: Road Infrastructure Safety Protection - Core Research and development for Road Safety in Europe; Increasing Safety and Reliability of Secondary Roads for a Sustainable Surface Transport, 135.
- ROQUE, C e JALAYER, M (2018). "Improving roadside design policies for safety enhancement using hazardbased duration modeling". Accident Analysis & Prevention, Volume 120, Pages 165-173, ISSN 0001-4575. <https://doi.org/10.1016/j.aap.2018.08.008>

- SAYED, T. and LEUR, P. d. (2000). *Developing a systematic framework for proactive road safety planning*. In: 13th ICTCT workshop - Traffic Safety Measures, Corfu - Greece.
- SECO, Álvaro; FERREIRA, Sara; SILVA, Ana; COSTA, Américo (2008). Manual do Planeamento de Acessibilidades e Transportes, 11 – Segurança Rodoviária
- SILVA, Philippe Barbosa (2019) - *Modelos de previsão de acidentes em rodovia brasileira de pista dupla*. Brasília: Departamento de Engenharia Civil e Ambiental, Universidade de Brasília. Tese de Doutorado.
- SINAY, M. C. F. & TAMAYO, A. S. (2005) Segurança Viária: Uma Visão Sistêmica. Anais do III Rio de Transportes, Rio de Janeiro, Brasil.
- SOULEYRETTE, R. R., HAAS, R. P. & MAZE, T. H. (2007) *Validation and implication of segmentation on Empirical Bayes for highway safety studies*. WIT Transactions on Biomedicine and Health, v. 11, p. 85-94.
- SOUZA, Suellen (2021) - *Identificação e análise de fatores de risco em autoestrada: aplicação de modelos estatísticos*. Dissertação de Mestrado, Faculdade de Engenharia da Universidade do Porto.
- WASHINGTON, S., KARLAFTIS, M. e MANNERING, F. (2003). *Statistical and econometric methods for transportation data analysis*.
- WRIGHT, C.C, BOYLE A.J., (1987) – *Road Accident causation and engineering treatment: a review of some current issues – Traffic Engineering and control*, pp 475-483, September
- ZENG, Q., HUANG, H., PEI, X. & WONG, S. C. (2016b) *Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks*. Analytic Methods in Accident Research, v. 10, p. 12–25.
- [https://peltarion.com/knowledge-center/documentation/modeling-view/build-an-ai-model/loss-functions/mean-squared-logarithmic-error-\(msle\)](https://peltarion.com/knowledge-center/documentation/modeling-view/build-an-ai-model/loss-functions/mean-squared-logarithmic-error-(msle))

ANEXO A

TABELA MATRIZ DE CORRELAÇÃO

Cor.Mat.1	LN_TMDSA	LN_COMPR	CARACT01	CARACT02	CARACT03	RAIO_C01	RAIO_C02	RAIO_C03	RAIO_C04	RAIO_C05	LARBEX01	LARBEX02	LARBEX03	OUTRVI01	OUTRVI02	OUTRVI03	OUTRVI05	OUTRVI06	N_VIAS01	N_VIAS02	N_VIAS03	LIM_VE01	LIM_VE02	LIM_VE03	LARG_NJ	TIPO_SEP	LARG_BIN	PERC_PES
LN_TMDSA	1.00000	-0.07462																										
LN_COMPR	-0.07462	1.00000																										
CARACT01	-0.04886	0.26048	1.00000																									
CARACT02	-0.22361	-0.25306	-0.30966	1.00000																								
CARACT03	0.23706	.00976	-0.54578	-0.62774	1.00000																							
RAIO_C01	-0.04886	.26048	1.00000	-0.30966	-0.54578	1.00000																						
RAIO_C02	-0.43693	-0.23634	-0.39223	-0.78949	-0.37462	-0.39223	1.00000																					
RAIO_C03	-0.06922	.02411	-0.15385	-0.17695	.28188	-0.15385	-0.22413	1.00000																				
RAIO_C04	0.31435	-0.05344	-0.14313	-0.16462	.26224	-0.14313	-0.20851	-0.08179	1.00000																			
RAIO_C05	0.37822	.03218	-0.31774	-0.36546	-0.58219	-0.31774	-0.46291	-0.18157	-0.16892	1.00000																		
LARBEX01	0.07506	-0.09340	-0.07451	-0.08569	.13651	-0.07451	-0.10855	-0.04257	-0.03961	.23448	1.00000																	
LARBEX02	-0.08433	.09881	.41890	-0.02828	-0.31802	.41890	-0.12148	-0.13500	-0.09841	-0.06538	1.00000																	
LARBEX03	0.05394	-0.06126	-0.37455	.05770	.25578	-0.37455	.15511	.14450	.13443	.01047	-0.29464	-0.93430	1.00000															
OUTRVI01	-0.14491	.04758	-0.16714	.05176	.09122	-0.16714	.07236	.17581	.04946	-0.06079	-0.10385	-0.32931	.35246	1.00000														
OUTRVI02	0.10291	-0.04537	-0.09905	.06748	.02162	-0.09905	.10102	.01981	.04213	-0.05455	-0.07675	-0.24338	.26049	-0.38659	1.00000													
OUTRVI03	0.08830	-0.06196	-0.03962	-0.04294	.07028	-0.03962	.00000	-0.06934	.04213	.05455	-0.07675	-0.24338	.26049	-0.38659	-0.28571	1.00000												
OUTRVI05	-0.06455	.07937	.26153	.05967	-0.26669	.26153	-0.02340	-0.11012	-0.10244	-0.08844	-0.05333	-0.81567	-0.76208	-0.26860	-0.19852	-0.19852	1.00000											
OUTRVI06	0.03613	-0.01348	.17971	-0.18872	.01919	.17971	-0.23905	-0.09376	-0.08723	.20082	.45408	.32185	-0.47044	-0.22871	-0.16903	-0.16903	-0.11744	1.00000										
N_VIAS01	-0.34466	-0.04010	-0.10647	.11068	-0.01038	-0.10647	.27145	-0.06084	-0.05660	-0.12566	-0.02946	-0.09343	.10000	.28372	-0.10968	-0.10968	-0.07621	-0.06489	1.00000									
N_VIAS02	0.13190	.02709	-0.04966	-0.01447	.05340	-0.04966	-0.11915	.08179	.07609	.08044	.03961	.12560	-0.13443	-0.21542	.14744	-0.04213	.10244	.08723	-0.74390	1.00000								
N_VIAS03	0.19866	.00555	.19655	-0.10550	-0.06794	.19655	-0.13363	-0.05241	-0.04876	.02406	-0.02538	-0.08049	.08615	-0.00376	-0.09449	.18898	-0.06565	-0.05590	-0.03627	-0.64087	1.00000							
LIM_VE01	0.30986	.05158	.10599	-0.13764	.03452	.10599	-0.17434	-0.06838	.47622	-0.14123	-0.03312	-0.10501	.11240	.02747	-0.01233	.09862	-0.08565	-0.07293	-0.04732	.06362	-0.04077	1.00000						
LIM_VE02	0.37830	-0.01100	.01774	-0.23203	.18996	.01774	-0.29390	-0.11528	.24277	.23198	-0.05583	-0.17703	.18948	.03372	.00799	.07993	-0.14440	-0.01892	-0.07978	-0.24277	.45468	-0.08967	1.00000					
LIM_VE03	-0.50719	-0.02833	-0.07571	.28133	-0.18594	-0.07571	.35635	.13977	-0.48297	-0.12295	.06769	.21464	-0.22973	-0.04512	.00000	-0.12599	.17508	.05797	.09673	.17647	-0.37500	-0.48925	-0.82476	1.00000				
LARG_NJ	-0.05394	.06126	.37455	-0.05770	-0.25578	.37455	-0.15511	-0.14450	-0.13443	-0.01047	.29464	.93430	1.00000	-0.35246	-0.26049	-0.26049	.76208	.47044	-0.10000	.13443	-0.08615	-0.11240	-0.18948	.22973	1.00000			
TIPO_SEP	-0.05394	.06126	.37455	-0.05770	-0.25578	.37455	-0.15511	-0.14450	-0.13443	-0.01047	.29464	.93430	1.00000	-0.35246	-0.26049	-0.26049	.76208	.47044	-0.10000	.13443	-0.08615	-0.11240	-0.18948	.22973	1.00000	1.00000		
LARG_BIN	0.16137	.08444	-0.07451	-0.08569	.13651	-0.07451	-0.10855	-0.04257	.52056	-0.08793	-0.02062	-0.06538	.06998	.04734	-0.07675	.09594	-0.05333	-0.04541	-0.02946	.03961	-0.02538	-0.03312	.36932	-0.30460	-0.06998	-0.06998	1.00000	
LN_PERC	-0.93653	0.06173	0.06541	0.22865	-0.25505	0.06541	0.43032	0.07920	-0.31949	-0.38941	-0.08551	0.10856	-0.07341	0.11593	-0.07621	-0.09681	0.08296	-0.03546	0.31315	-0.09528	-0.21723	-0.26417	-0.38600	0.48799	0.07341	0.07341	-0.17645	1.00000

ANEXO B

RESULTADOS DOS MODELOS SOFTWARE NLOGIT VERSÃO 5

ANEXO B.1 – MODELO DE FREQUÊNCIA DE ACIDENTES TOTAIS - A44 (15-20)

Resultado Poisson

```

-----
Poisson Regression
Dependent variable      ACID_TOT
Log likelihood function  -437.58931
Restricted log likelihood -528.03936
Chi squared [ 12 d.f.]  180.90009
Significance level      .00000
McFadden Pseudo R-squared .1712941
Estimation based on N = 594, K = 13
Inf.Cr.AIC = 901.2 AIC/N = 1.517
Model estimated: May 11, 2021, 12:30:01
Chi- squared = 631.27164 RsqP= .2987
G - squared = 502.15071 RsqD= .2648
Overdispersion tests: g=mu(i) : 1.775
Overdispersion tests: g=mu(i)^2: 1.818
-----+-----
|          Standard      Prob.  95% Confidence
ACID_TOT| Coefficient    Error   z   |z|>Z*   Interval
-----+-----
Constant| -19.8978***    4.50155  -4.42 .0000  -28.7207 -11.0749
LN_TMDSA|  1.49380***    .43331   3.45 .0006   .64453  2.34306
LN_COMPR|  .83774***     .11832   7.08 .0000   .60584  1.06963
RAIO_C02| -.36526*       .21405  -1.71 .0879  -.78479 .05426
RAIO_C03| -1.25997***    .39543  -3.19 .0014  -2.03499 -.48495
RAIO_C04| -.64394**      .26771  -2.41 .0162  -1.16864 -.11924
RAIO_C05| -.64708***     .20773  -3.11 .0018  -1.05423 -.23993
OUTRVI02| .51057***      .18863   2.71 .0068   .14086 .88028
OUTRVI03| .40103**       .18831   2.13 .0332   .03194 .77012
OUTRVI05| .02828         .25193   .11 .9106  -.46550 .52206
OUTRVI06| .61073**       .27241   2.24 .0250   .07682 1.14465
LIM_VE01| .62594**       .28309   2.21 .0270   .07110 1.18078
LIM_VE02| -.87726***     .28131  -3.12 .0018  -1.42862 -.32591
-----+-----

```

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

Normal exit: 20 iterations. Status=0, F= 435.6417

Resultado Binomial Negativo

Negative Binomial Regression

Dependent variable ACID_TOT
 Log likelihood function -435.64169
 Restricted log likelihood -437.58931
 Chi squared [1 d.f.] 3.89525
 Significance level .04842
 McFadden Pseudo R-squared .0044508
 Estimation based on N = 594, K = 14
 Inf.Cr.AIC = 899.3 AIC/N = 1.514
 Model estimated: May 11, 2021, 12:30:01
 NegBin form 2; Psi(i) = theta
 Tests of Model Restrictions on Neg.Bin.
 Model Logl ChiSquared[df]
 Poisson(b=0) -528.04 ***** [**]
 Poisson -437.59 180.9 [12]
 Negative Bin. -435.64 3.9 [1]

	Standard	Prob.	95% Confidence		
ACID_TOT	Coefficient	Error z	z >Z*	Interval	
Constant	-19.6222***	4.91063	-4.00 .0001	-29.2469	-9.9976
LN_TMDDA	1.45549***	.47483	3.07 .0022	.52483	2.38615
LN_COMPR	.85576***	.13700	6.25 .0000	.58724	1.12428
RAIO_C02	-.35144	.23658	-1.49 .1374	-.81513	.11225
RAIO_C03	-1.24562**	.48457	-2.57 .0102	-2.19537	-.29588
RAIO_C04	-.64918*	.36374	-1.78 .0743	-1.36210	.06374
RAIO_C05	-.65405***	.24558	-2.66 .0077	-1.13538	-.17272
OUTRVI02	.53043**	.21095	2.51 .0119	.11698	.94388
OUTRVI03	.46591**	.21286	2.19 .0286	.04872	.88311
OUTRVI05	.05590	.28455	.20 .8443	-.50180	.61360
OUTRVI06	.64925**	.30911	2.10 .0357	.04341	1.25508
LIM_VE01	.70923*	.36871	1.92 .0544	-.01342	1.43189
LIM_VE02	-.83811***	.29349	-2.86 .0043	-1.41335	-.26287
Dispersion parameter for count data model					
Alpha	.24987	.16505	1.51 .1301	-.07362	.57336

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

ANEXO B.2 – MODELO DE FREQUÊNCIA DE ACIDENTES COM VÍTIMAS – A44 (15-20)

Resultado Poisson

Poisson Regression

Dependent variable ACID_VIT
 Log likelihood function -250.67689
 Restricted log likelihood -295.98287
 Chi squared [8 d.f.] 90.61196
 Significance level .00000
 McFadden Pseudo R-squared .1530696
 Estimation based on N = 594, K = 9
 Inf.Cr.AIC = 519.4 AIC/N = .874
 Model estimated: May 11, 2021, 12:30:13
 Chi- squared = 663.37054 RsqP= .1134
 G - squared = 324.26875 RsqD= .2184
 Overdispersion tests: $g=\mu(i)$: 1.024
 Overdispersion tests: $g=\mu(i)^2$: .230

	Standard	Prob.	95% Confidence		
ACID_VIT	Coefficient	Error	z	z >Z*	Interval
Constant	-27.1364***	5.60506	-4.84	.0000	-38.1221 -16.1507
LN_TMDSA	2.23057***	.52486	4.25	.0000	1.20185 3.25928
LN_COMPRA	.74560***	.16874	4.42	.0000	.41487 1.07633
TIPO_SEP	-.61111**	.28502	-2.14	.0320	-1.16973 -.05248
RAIO_C02	-.97853***	.29396	-3.33	.0009	-1.55469 -.40238
RAIO_C03	-1.93441***	.60747	-3.18	.0015	-3.12503 -.74378
RAIO_C04	-1.25452***	.42400	-2.96	.0031	-2.08554 -.42350
RAIO_C05	-1.13621***	.26805	-4.24	.0000	-1.66158 -.61085
LIM_VE02	-1.55247***	.43363	-3.58	.0003	-2.40237 -.70257

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

Resultado Binomial Negativo

Negative Binomial Regression

Dependent variable ACID_VIT
 Log likelihood function -250.56638
 Restricted log likelihood -250.67689
 Chi squared [1 d.f.] .22104
 Significance level .63825
 McFadden Pseudo R-squared .0004409
 Estimation based on N = 594, K = 10
 Inf.Cr.AIC = 521.1 AIC/N = .877
 Model estimated: May 11, 2021, 12:30:13
 NegBin form 2; Psi(i) = theta
 Tests of Model Restrictions on Neg.Bin.
 Model Logl ChiSquared[df]
 Poisson(b=0) -295.98 ***** [**]
 Poisson -250.68 90.6 [8]
 Negative Bin. -250.57 .2 [1]

	Standard	Prob.	95% Confidence		
ACID_VIT Coefficient	Error	z	z >Z*	Interval	
Constant	-27.0953***	6.65733	-4.07 .0000	-40.1434	-14.0472
LN_TMDA	2.22465***	.62226	3.58 .0004	1.00504	3.44426
LN_COMPR	.75025***	.17875	4.20 .0000	.39991	1.10059
TIPO_SEP	-.61001*	.33485	-1.82 .0685	-1.26630	.04627
RAIO_C02	-.97767***	.34315	-2.85 .0044	-1.65023	-.30511
RAIO_C03	-1.93711***	.66960	-2.89 .0038	-3.24951	-.62471
RAIO_C04	-1.24848**	.50272	-2.48 .0130	-2.23380	-.26317
RAIO_C05	-1.14308***	.31256	-3.66 .0003	-1.75569	-.53047
LIM_VE02	-1.54694***	.47744	-3.24 .0012	-2.48270	-.61119
Dispersion parameter for count data model					
Alpha	.11965	.25783	.46 .6426	-.38568	.62499

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

ANEXO B.3 – MODELO DE FREQUÊNCIA DE ACIDENTES TOTAIS– A44 (15-18)

Resultado Poisson

Poisson Regression

Dependent variable ACID_TOT
 Log likelihood function -362.72871
 Restricted log likelihood -435.69631
 Chi squared [12 d.f.] 145.93521
 Significance level .00000
 McFadden Pseudo R-squared .1674735
 Estimation based on N = 495, K = 13
 Inf.Cr.AIC = 751.5 AIC/N = 1.518
 Model estimated: May 21, 2021, 17:25:23
 Chi- squared = 510.72699 RsqP= .3026
 G - squared = 417.05972 RsqD= .2592
 Overdispersion tests: $g=\mu(i)$: 1.855
 Overdispersion tests: $g=\mu(i)^2$: 1.893

	Standard	Prob.	95% Confidence		
ACID_TOT	Coefficient	Error	z	z >Z*	Interval
Constant	-17.4057***	4.84352	-3.59	.0003	-26.8988 -7.9126
LN_TMDSA	1.23663***	.46703	2.65	.0081	.32127 2.15198
LN_COMPR	.85254***	.13149	6.48	.0000	.59483 1.11025
RAIO_C02	-.32830	.23566	-1.39	.1636	-.79019 .13359
RAIO_C03	-1.29438***	.45205	-2.86	.0042	-2.18038 -.40839
RAIO_C04	-.45377	.28842	-1.57	.1157	-1.01907 .11153
RAIO_C05	-.58806***	.22626	-2.60	.0093	-1.03153 -.14460
OUTRVI02	.57197***	.20867	2.74	.0061	.16297 .98096
OUTRVI03	.49626**	.20867	2.38	.0174	.08728 .90524
OUTRVI05	.00995	.28448	.03	.9721	-.54763 .56753
OUTRVI06	.80220***	.29370	2.73	.0063	.22656 1.37784
LIM_VE01	.65442**	.31280	2.09	.0364	.04134 1.26751
LIM_VE02	-.74012**	.29923	-2.47	.0134	-1.32660 -.15363

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

ANEXO B.4 – MODELO DE FREQUÊNCIA DE ACIDENTES COM VÍTIMAS – A44 (15-18)

Resultado Poisson

Poisson Regression

Dependent variable ACID_VIT
 Log likelihood function -208.52564
 Restricted log likelihood -244.49006
 Chi squared [8 d.f.] 71.92885
 Significance level .00000
 McFadden Pseudo R-squared .1470997
 Estimation based on N = 495, K = 9
 Inf.Cr.AIC = 435.1 AIC/N = .879
 Model estimated: May 21, 2021, 17:29:39
 Chi-squared = 553.61604 RsqP= .1116
 G - squared = 270.42105 RsqD= .2101
 Overdispersion tests: $g=\mu(i)$: 1.023
 Overdispersion tests: $g=\mu(i)^2$: .295

	Standard	Prob.	95% Confidence		
ACID_VIT	Coefficient	Error	z	z >Z*	Interval
Constant	-24.2411***	6.00077	-4.04	.0001	-36.0024 -12.4798
LN_TMDA	1.95599***	.56241	3.48	.0005	.85368 3.05830
LN_COMPR	.73189***	.18422	3.97	.0001	.37083 1.09296
TIPO_SEP	-.64984**	.31525	-2.06	.0393	-1.26772 -.03195
RAIO_C02	-1.09446***	.33326	-3.28	.0010	-1.74763 -.44129
RAIO_C03	-2.12025***	.73752	-2.87	.0040	-3.56577 -.67473
RAIO_C04	-1.01025**	.43490	-2.32	.0202	-1.86264 -.15786
RAIO_C05	-.99307***	.28803	-3.45	.0006	-1.55759 -.42854
LIM_VE02	-1.33768***	.43668	-3.06	.0022	-2.19356 -.48180

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

ANEXO B.5 – MODELO DE FREQUÊNCIA DE ACIDENTES TOTAIS– A29 (15-17)

Resultado Poisson

```
-----
Poisson Regression
Dependent variable          ACID_TOT
Log likelihood function      -1262,7
Restricted log likelihood    -1435,53
Chi squared [ 8 d.f.]      345,6612
Significance level          0
McFadden Pseudo R-squared   0,120395
Estimation based on N = 2   800, K = 9
Inf.Cr.AIC = 2543.4 AIC/    N = .908
Model estimated: Mar 03, 20  21, 11:05:33
Chi- squared = 3488.95816    RsqP= .2223
G - squared = 1705.58094     RsqD= .1685
Overdispersion tests: g=mu(  i) : 2.565
```

Overdispersion tests: g=mu(i)^2: 3.017

```
-----+-----
|          Standard      Prob.  95% Confidence
ACID_TOT| Coefficient   Error   z  |z|>Z*   Interval
-----+-----
Constant| -13.2377***   1.99042  -6.65 .0000  -17.1389 -9.3366
LN_TMADA|  .75966***    .19676   3.86 .0001   .37402  1.14530
LN_COMPR|  .94735***    .06500  14.57 .0000   .81994  1.07475
ANO_CAT|  -.06537**    .03140  -2.08 .0374   -.12692 -.00382
N_VIAS|  .75727***    .25230   3.00 .0027   .26276  1.25177
LARG_VIA| -.93154***    .27145  -3.43 .0006   -1.46357 -.39951
RAIO_CAT|  .23800**     .09728   2.45 .0144   .04733  .42866
L_BER1|  -.22819     .17845  -1.28 .2010   -.57794 .12156
VELOC_02|  .30988**     .15647   1.98 .0477   .00321  .61655
-----+-----
```

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.