



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO ACADÊMICO DO AGRESTE  
NÚCLEO DE TECNOLOGIA  
CURSO DE GRADUAÇÃO EM ENGENHARIA CIVIL

LUCAS HENRIQUE DA CRUZ FERREIRA

**ANÁLISE DOS FATORES CONTRIBUINTE NA OCORRÊNCIA DE ACIDENTES  
EM AUTOESTRADAS DA REGIÃO NORTE DE PORTUGAL**

Caruaru

2024

LUCAS HENRIQUE DA CRUZ FERREIRA

**ANÁLISE DOS FATORES CONTRIBUINTE NA OCORRÊNCIA DE ACIDENTES  
EM AUTOESTRADAS DA REGIÃO NORTE DE PORTUGAL**

Trabalho de Conclusão de Curso apresentado à  
Coordenação do Curso de Engenharia Civil do  
Campus Agreste da Universidade Federal de  
Pernambuco – UFPE, na modalidade de artigo  
científico, como requisito parcial para obtenção  
do grau de Bacharel em Engenharia Civil.

**Área de concentração:** Transportes

**Orientadora:** Profa. Dra. Jocilene Otilia da Costa

Caruaru

2024

## AGRADECIMENTOS

Aos meus avós: Manoel e Prescila, por sempre me apoiarem e incentivarem a seguir os meus desejos, por me ensinarem princípios que eu levarei para o resto da minha vida a transmitirei para futuras gerações.

À minha mãe, Adilma, por ter me ensinado tudo que sou e me apoiado independentemente a situação. Por ter me mostrado o que é o amor, e por sempre ter me incentivado a lutar pelos meus sonhos.

Ào Rômulo, por ter sempre me apoiado em situações complicadas, e sempre me incentivando a ser melhor e alcançar patamares maiores em minha vida pessoal e profissional.

Aos professores Dannúbia Pires e Edevaldo Miguel, por aceitarem participar da banca de avaliação deste trabalho e contribuírem com sugestões para o seu aperfeiçoamento.

À minha orientadora e amiga, Jocilene Costa, por ter me acompanhado nesta fase acadêmica, por todas as conversas e conselhos em situações de dificuldade. Por ser uma exímia profissional e repassar ensinamentos importantes a todos os alunos que certamente os acompanharão pela sua vida pessoal e profissional.

## **Análise dos fatores contribuintes na ocorrência de acidentes em autoestradas da região norte de Portugal**

### **Analysis of contributing factors in the occurrence of accidents on motorways in the northern region of Portugal**

**Lucas Henrique da Cruz Ferreira<sup>1</sup>**

---

#### **RESUMO**

A modelagem de dados para previsões de acidentes em rodovias tem tomado popularidade nos últimos 20 anos com a finalidade de melhorar a segurança nas principais rotas dos países, mas também como forma de evitar grandes prejuízos financeiros em decorrência das sinistralidades. Reconhecer estes fatores que contribuem tanto positivamente quanto negativamente para a ocorrência destes acidentes se torna um ponto crucial para o início da criação de programas de prevenção e controle destas situações. O estudo atual utilizou os modelos lineares generalizados mistos com as distribuições de Poisson e Binomial negativa para identificar estes fatores contribuintes em trechos de 500 metros antes e após a sinistralidade considerando as características do pavimento, mas também os fatores relacionados as características geométricas e variáveis complementares. Os acidentes de viação representam uma enorme carga social e econômica que impacta na sociedade, o aumento da procura dos meios de transporte e desenvolvimento das cidades tem resultado em um visível aumento no número de acidentes. Este estudo visa a possibilidade de maiores entendimentos sobre a origem destes acidentes, como também a viabilidade de elaboração futura de programas para a prevenção destes ocorridos em virtude dos fatores identificados em decorrência da modelagem elaborada. Por fim, o estudo indicou que para as modelagens elaboradas tanto com a função de Poisson quanto para a função da Binomial negativa, as variáveis relacionadas ao raio da curva, parâmetro A, alinhamentos retos, extensão da curva circular e Clotóide são parâmetros significativos para a previsão de acidentes, apresentando um erro quadrático médio aproximado para as duas funções.

**Palavras-chave:** Acidentes; autoestradas; modelos de previsão; modelos de zeros inflacionados.

<sup>1</sup>Graduando em Engenharia Civil pela Universidade Federal de Pernambuco. E-mail: lucas.ferreira@ufpe.br

---

## ABSTRACT

Data modeling for highway accident predictions has gained popularity over the last 20 years. Its purpose is twofold: to enhance safety on major roadways in countries and to prevent significant financial losses due to accidents. Recognizing factors that contribute both positively and negatively to accident occurrences is crucial for developing prevention and control programs. The current study employed mixed-effects generalized linear models using Poisson and negative binomial distributions to identify contributing factors within 500-meter segments before and after accidents. These factors include pavement characteristics, geometric features, and complementary variables. Road accidents impose a substantial social and economic burden on society. The increasing demand for transportation and urban development has led to a visible rise in accident numbers. This study aims to provide deeper insights into the origins of these accidents and explore the feasibility of future prevention programs based on the identified factors from the developed model. Finally, the study indicated that for the models developed using both the Poisson function and the negative binomial function, the variables related to curve radius, parameter A, straight alignments, circular curve extension, and clothoid are significant parameters for accident prediction, with an approximate mean squared error for both functions.

**Keywords:** accidents; highways; prediction models; inflated zero models.

---

**DATA DE APROVAÇÃO:** 17 de maio de 2024.

---

## 1 INTRODUÇÃO

Sendo um dos sistemas de transporte mais utilizados em Portugal, o sistema rodoviário representa uma parcela nos fatores que influenciam socialmente e economicamente a sociedade. Em 2020, Portugal possuía cerca de 14.325 km de extensão em sua rede nacional rodoviária, sendo destes 3.065 km sendo autoestradas que percorrem todo o continente correspondendo a cerca de 21,40% de sua malha viária (INE, 2020).

Todos os anos, aproximadamente 1,3 milhões de pessoas no mundo morrem decorrente de acidentes de trânsito. Entre 20 e 50 milhões de pessoas sofrem lesões não fatais, sendo a maioria resultando em algum tipo de deficiência (WHO, 2022). Somente em Portugal, no ano de 2021,

houve 30.691 acidentes com vítimas, dos quais resultaram 401 vítimas mortais, 2.297 feridos graves e 35.877 feridos leves. Face ao ano de 2020, houve uma diminuição de 0,7% de vítimas mortais, e um aumento de 10,7%, 15,1% e 11,9% respectivamente (ANSR, 2022).

Em Portugal, no ano de 2021, registrou-se um aumento no tráfego médio diário nas autoestradas no valor de 17.349 veículos/dia (IMT-IP, 2022). No mesmo ano houve um índice de gravidade, vítimas mortais por cada 100 acidentes, para autoestradas de 2,12 vítimas mortais que representa uma redução de 38,4%, e um aumento de 9,3% de acidentes nestas vias comparado ao ano anterior (ANSR, 2021).

As análises de segurança de sistemas de transportes contêm em sua grande maioria a análise do risco de acidentes em seus indicadores. O risco é um dos principais fatores de comparação para o indicador de incidência de acidentes. Sendo a definição mais tradicional de risco, a relação entre números de acidentes e a quantidade de exposições em seu modal de transporte (Lord, D., 2002).

Segundo o IPEA (2008) existem inúmeros fatores que influenciam a ocorrência de acidentes em rodovias, como condições geométricas da via, clima desfavorável para o tráfego de veículos, e principalmente condições humanas ou decorrentes do mesmo como comportamento inadequado, e frota de veículos em estado não ideal para uso. Todos estes fatores formam um conjunto de indicadores que causam como resultado o acontecimento de acidentes nas estradas.

O acidente rodoviário ocorre quando se verifica uma ruptura entre os diversos componentes do trânsito, ou seja, entre o usuário, o veículo e o ambiente. No geral, é aceito que a ocorrência de um acidente inclui três fases, nomeadamente a fase de percepção, a fase de decisão e a fase de conflito (Mário Talaia, 2012).

TBR (2004) apresenta que a análise dos problemas de segurança viária não deve ser restrita a somente à análise da soma total e severidades dos acidentes de trânsito. É o conjunto de condições e fatores interligados que proporcionam esta segurança.

Apesar da tendência em subestimar os efeitos das características e das condições da via sobre a ocorrência de acidentes de trânsito, o sistema viário, no seu conjunto, cria situações que podem induzir os motoristas a cometerem erros de percepção ou de reação e, conseqüentemente, propícias para a ocorrência de acidentes (Maia, 1995).

É de conhecimento comum que a principal tipologia de causadores de acidentes é o fator humano, sendo veemente utilizada a intervenção viária como uma tentativa de reduzir a incidência de acidentes com este fato. Nodari e Lindau (2001) apud Sampedro (2010) apontam que ambientes viários mais complexos resultam em mais habilidades requeridas do motorista,

resultando em mais incidência de acidentes. Aponta também que é mais eficiente aplicar melhorias no fator via do que em medidas de educação ou segurança, já que ela permite uma grande melhoria rápida da gravidade e incidência de acidentes de trânsito.

Caso não utilizado o método correto, os dados para o estudo de acidentes em estradas podem comprometer estatisticamente a análise feita. Para isso, existem uma série de modelos de análise que podem atender melhor os dados obtidos, tendo como ponto de partida o modelo de Regressão de Poisson (Lord D., 2010).

Por não ser possível adotar um único modelo que englobe a rede rodoviária, com o passar das décadas foram adaptados modelos de previsão a fim de melhor se adaptar ao segmento estudado (Costa J., 2018). Como os modelos de Markov Switching e Latent-Class para lidar com a heterogeneidade não observada, além do surgimento de modelos multivariados incluindo efeitos do tempo e/ou espaço, e questões de endogeneidade (Mannering, 2010).

Por conta da diversidade de metodologias para a pesquisa, é possível estudar diferentes dados de acidentes de trânsito, um dos mais utilizados é o modelo binomial negativo (Poisson-Gamma), por lidar com dados super dispersos (Mannering, 2014). E alguns pesquisadores como Lord (2007) ao trabalhar com dados com muitas observações sem colisões consideraram Poisson inflado de zero e regressões binomiais negativas para a sua análise a fim de explicar a preponderância de zeros.

Dentro do contexto dissertado, o presente trabalho busca identificar os fatores que causam acidentes nas autoestradas da região norte de Portugal. A partir de uma análise estatística de dados levantados não somente de sinistralidades, mas também de indicadores geométricos das vias. O trabalho busca apresentar uma relação entre os indicadores utilizados e a incidência de acidentes nas autoestradas analisadas, de modo que seja possível propor estudos de melhoria de segurança viária nestas regiões.

### **1.1 Fatores contribuintes na ocorrência de acidentes viários**

Os fatores contribuintes conseguem identificar a origem do sinistro e possibilitam a elaboração de prevenção a estes acidentes. Eles englobam diferentes tipologias, como ações, falhas ou condições que levam diretamente ao acidente (Sampaio, 2018), organizados na tabela 1.

**Tabela 1** – Fatores contribuintes para acidentes

<b>Fator</b>	<b>Índice</b>
Fator Veículo	Idade do veículo
	Revisões periódicas
	Falha mecânica
	Caracterização do veículo (marca)
	Sistemas de navegação incorporado
	Sistema de ABS
Fator Via e Ambiente	Estado do piso
	Desenho da estrada
	Proteções da via
	Infraestrutura rodoviária
	Condições meteorológicas
Deficiente sinalização	
Fator Humano	Sexo
	Atitudes
	Idade
	Fadiga
	Estado de espírito
	Cultura
	Ética
	Crenças, valores e opiniões
	Comportamentos de risco
Experiência de condução	

Fonte: Adaptado de Sampaio (2018)

## 1.2 Modelos lineares generalizados mistos

Sendo uma extensão dos modelos lineares generalizados, os modelos lineares generalizados mistos levam em consideração os efeitos aleatórios no preditor linear, permitindo a modelagem da estrutura de correlação das observações de um mesmo indivíduo (Souza, 2022).

Inferindo sobre o indivíduo, os modelos lineares generalizados descrevem as relações de cada um com as covariáveis de interesse, como descrito em Cabral e Gonçalves (2011).

### 1.2.1 Estrutura do modelo linear generalizado misto

Para a especificação dos modelos lineares generalizados mistos, deve-se iniciar com o condicionamento da variável resposta aos efeitos aleatórios, seguindo uma distribuição ligada à família exponencial e com função densidade.

$$f(y_{it}|b_i) = \exp\left\{\frac{\omega_{it}}{\phi}(y_{it}\theta_{it} - b(\theta_{it}) + c(y_{it}, \phi))\right\}$$

Retomando o valor médio e a variância condicional, pode-se expressar por:

$$E(Y_{it}|b_i) = \mu_{it} = b'(\theta_{it})$$

$$Var(Y_{it}|b_i) = \sigma_{it} = b''(\theta_{it}) \frac{\phi}{\omega_{it}}$$

Por ser uma extensão aos modelos lineares generalizados, é possível realizar a seguinte definição:

$$g(\mu_{it}) = x_{it}^T \beta + z_{it}^T b_i$$

$$\sigma_{it} = V(\mu_{it}) \frac{\phi}{\omega_{it}}$$

Onde:

$g$  = é uma função de ligação;

$V$  = é a função de variância;

$\phi$  = é o parâmetro de dispersão;

$\omega_{it}$  = é uma constante conhecida atribuída às observações da variável resposta;

$y_{it}$  = representa o vetor das variáveis resposta para o indivíduo  $i$ ;

$\beta$  = é o vetor dos parâmetros de regressão desconhecidos;

$x_i$  = é o vetor de covariáveis associado a um vetor de efeitos aleatórios;

$b_i$  = é o vetor de efeitos aleatórios.

A partir do momento em que são fixados os efeitos aleatórios do modelo, as variáveis resposta se definem independentes.

### 1.2.2 Modelos para dados de contagem com excesso de zeros

Tendo em vista que os dados utilizados têm grande presença de zeros e sobredispersão, o modelo adotado para realizar as modelagens serão os modelos com zeros inflacionados, a fim de contornar o problema inicial.

Esses modelos atuam como uma mistura de duas distribuições distintas com dois processos paralelos: (i) um processo que trata a problemática da grande quantidade de zeros, modelado por uma massa pontual em zero; e (ii) um processo que inclui as contagens nulas modelado por uma Binomial Negativa ou Poisson. Entende-se que o primeiro processo ocorre com uma probabilidade  $\pi_i$  enquanto o segundo processo ocorre com uma probabilidade  $1 - \pi_i$  (Silva, 2015). Em resumo, para cada variável definida ocorrem dois processos distintos para cada modelo observado.

### 1.2.3 Modelo de Poisson Inflacionado de Zeros

Comumente utilizado em modelagem com dados que contém excesso de zeros, o modelo de Poisson Inflacionados de Zeros (ZIP) foi inicialmente proposto por Lambert (1992). Para a amostra com  $n$  observações, a variável resposta é dada por  $Y_i$ , onde  $i = 1, 2, \dots, n$  e das covariáveis é dada por  $X = (X_2, \dots, X_p)$ . Assumindo que a variável  $y_i, \dots, y_n$  resultados da variável  $Y_i$ , o modelo é dado por:

$$P(Y_i) = y_i | x_i = \begin{cases} \pi_i + (1 - \pi_i)e^{-\mu}, & y_i = 0 \\ (1 - \pi_i) \frac{e^{-\mu}}{y_i!}, & y_i > 0 \end{cases}$$

Onde a média e a variância da distribuição são obtidos por:

$$\begin{aligned} IE(y_i) &= (1 - \pi_i)\mu_i \\ Var(Y_i) &= \mu_i(1 - \pi_i)(1 + \pi_i\mu_i) \end{aligned}$$

É utilizado uma regressão logística com função de ligação para modelar a probabilidade de  $\pi$  e uma regressão de Poisson para modelar a média de  $\mu$  de uma variável de Poisson, ou seja,

$$\log(\mu_i) = X_i\beta$$

$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right) = G_i\gamma$$

Onde  $X_i$  e  $G_i$  são as matrizes de covariáveis, podendo ou não serem coincidentes.

#### 1.2.4 Modelo Binomial Negativo Inflacionado de Zeros

Analogamente ao modelo de Poisson Inflacionado de Zeros, o modelo Binomial Negativo Inflacionado de Zeros (ZINB) também é utilizado para dados com grande incidência de zeros, mas com ocorrência de sobredispersão. Proposto por Cheung (2002), embora similar ao ZIP, sua característica é que a variável  $Y_i$  segue uma distribuição binomial negativa de parâmetros  $(\mu_i, \alpha)$  com zeros inflacionados e que tem como probabilidade a seguinte expressão:

$$P(Y_i = y_i | x_i) = \begin{cases} \pi_i + (1 - \pi_i) \left(\frac{1}{1 + \alpha\mu_i}\right)^{\alpha^{-1}}, & y_i = 0 \\ (1 - \pi_i) \frac{r(y_i + \alpha^{-1})}{y_i! r(\alpha^{-1})} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i} \left(\frac{1}{1 + \alpha\mu_i}\right)^{\alpha^{-1}}, & y_i > 0 \end{cases}$$

Onde a constante  $\alpha$  é não negativa e parâmetro de heterogeneidade.

A média e a variância do modelo binomial negativo com zeros inflacionados são dados por:

$$E(y_i) = (1 - \pi_i)\mu_i$$

$$\text{Var}(Y_i) = \mu_i(1 - \pi_i)(1 + \pi_i\mu_i + \alpha\mu_i)$$

Como é possível ver, quando  $\alpha$  tende para zero a distribuição resume-se em uma distribuição de Poisson com Zeros Inflacionados (Zaihra, 2012).

#### 1.2.5 Ajuste do modelo de previsão de acidentes

É necessário um ajuste no modelo de previsão de acidentes e colisões para uma forma funcional. Utilizando as técnicas de modelação genéricas e de zeros inflacionados, com distribuição do erro binomial negativo e Poisson, é possível realizar o ajuste.

Considera-se a estimativa de acidentes para um dado elemento rodoviário, representada por

$Y_{mt}$ , ao longo do período de tempo  $t$ , é considerada a seguinte equação (Costa, Pereira, Freitas, & Jacques, 2012):

$$E(Y_m) = \alpha \left\{ \left[ \prod_i (a_{i,mt})^{\gamma_i} \right] e^{\sum_j \beta_j X_{j,mt}} \right\}$$

Onde:

$E(Y_m)$  = número esperado de acidentes para o elemento  $m$  no período de tempo  $t$ ;

$a_{i,mt}, X_{j,mt}$  = variáveis explicativas ( $i$  e  $j$ ) observadas no elemento  $m$  no tempo  $t$ ;

$a_t$  = parâmetro do modelo relacionado com o período de tempo  $t$ ;

$\gamma_i, \beta_j$  = parâmetros do modelo constante para todo  $t$

A análise dos modelos será feita de forma isolada, porém todos os grupos serão analisados aplicando os dois modelos, Poisson Inflacionado de Zeros e Binomial Negativo Inflacionado de Zeros.

### 1.3 Objetivos

#### 1.3.1 Objetivo geral

Este trabalho busca identificar os fatores contribuintes na ocorrência de acidentes em autoestradas da região norte de Portugal, utilizando metodologias de zeros inflacionados que para o estudo realizado será utilizado as modelagens de Poisson Inflacionado de Zeros e Binomial Negativo Inflacionado de Zeros

#### 1.3.2 Objetivos específicos

- Montagem da base de dados em relação as autoestradas;
- Tabular dados geométricos das áreas estudadas;
- Definir condições de contorno para as análises;
- Realizar as análises dos dados em relação às autoestradas.

---

## 2 METODOLOGIA

A metodologia aplicada para a análise dos fatores contribuintes em acidentes será dividida

em etapas a fim de obter maior entendimento nas análises e procedimentos utilizados.

Estas etapas também foram utilizadas por Costa (2013), organizadas na tabela 2, que desenvolveu um modelo de previsão de acidentes também para a região norte de Portugal, assim como o presente artigo.

**Tabela 2** – Etapas para modelagem de acidentes

<b>Agrupamento</b>	<b>Etapa</b>
Elementos viários	ETAPA 1 – Definição das autoestradas analisadas
	ETAPA 2 – Obtenção dos dados de geometria das autoestradas selecionadas
Amostra dos elementos viários	ETAPA 3 – Divisão dos trechos das Autoestradas em segmentos
	ETAPA 4 – Agrupamento dos trechos definidos
	ETAPA 5 – Definição das amostras
Base dados	ETAPA 6 – Levantamento dos dados dos acidentes nas Autoestradas
	ETAPA 7 – Levantamento dos índices das Autoestradas
	ETAPA 8 – Elaboração da base de dados do estudo
Modelagem	ETAPA 9 – Ajuste do modelo

Fonte: Autor (2024)

## **2.1 ETAPA 1 - Definição das autoestradas analisadas**

A definição das autoestradas utilizadas neste estudo se deu por meio da disponibilidade de dados apresentados pelas empresas gestoras de acordo com o relatório anual de verificação dos índices das vias que estão sob sua responsabilidade, e a fim de melhorar a segurança viária na região norte de Portugal tendo em vista que é a região mais populosa de Portugal, resultando em uma maior probabilidade de acidentes por via. A fim de evitar grandes impactos nos dados analisados, autoestradas que estavam com grandes quantidades de dados faltantes ou incoerentes com a realidade foram descartadas do cálculo.

## **2.2 ETAPA 2 - Obtenção dos dados de geometria das autoestradas selecionadas**

Os dados para esta análise foram fornecidos pelas empresas decorrente da análise anual das vias do período de 2014 a 2021 e associados a um ponto quilométrico de via. Estes dados vão ser usados posteriormente para constituir a base de dados das características das vias definidas no estudo. Assim os dados recolhidos preliminarmente das autoestradas foram:

- Índices de irregularidade longitudinal;
- Índices de rodeiras;
- Índices de profundidade de perfil;

- Índice de incidência de acidentes;
- Tráfego médio anual;

### 2.3 ETAPA 3 - Divisão dos trechos das Autoestradas em segmentos

Com os dados obtidos da etapa anterior, para uma melhor aplicabilidade do método de análise que será definido para o estudo, foi utilizado o conceito de ponto negro aplicado em Portugal, implicando em divisões de segmentos de 500 metros de comprimento por toda a via utilizada. Ao utilizar este conceito é possível identificar as zonas de acumulação de acidentes.

### 2.4 ETAPA 4 – Agrupamento dos trechos definidos

A partir dos trechos de 500 metros definidos, com a motivação de manter a modelagem mais precisa de acordo com os dados obtidos, se faz necessário a divisão dos dados em dois agrupamentos de regiões distintas.

Isso ocorre pois, de acordo com o observado, há regiões onde algumas características das vias não são medidas, inviabilizando uma modelagem de dados única que mantenha a precisão em um nível satisfatório devido à falta de dados em seletas regiões importantes para o estudo.

De frente a esta peculiaridade, se faz necessário a divisão dos dados de acordo com a concessão, conforme mostrado na Tabela 3:

**Tabela 3** – Distribuição dos agrupamentos

Concessão	Agrupamento
Costa de Prata - Grupo III	CPG3
Costa de Prata - Grupo IV	CPG4

### 2.5 ETAPA 5 – Definição das amostras

Com os trechos de 500 metros definidos, obteve-se um total de 1045 trechos para o estudo divididos em 2 agrupamentos, CPG3 e CPG4. Esta distribuição das amostras selecionadas está mais detalhada na Tabela 4, onde são apresentadas as localizações de cada amostra obtida:

**Tabela 4** – Distribuição dos trechos estudados

Agrupamento	Localização	Trechos
CPG3	Anjeja - Salreu	120

	Arada - Maceda	64
	Estarreja – Ovar Sul	144
	Obar Sul - Arada	71
	Salreu - Estarreja	80
	<hr/>	
	A29/A44 – Canelas	66
	A29/A44 – Gulpilhares	23
	A29/A44 - Francelos	16
	Canelas – IC2	42
	Cortegaça - Esmoriz	40
	Gulpilhares - Miramar	32
	Hospital – IP1	24
	IC2 - Hospital	12
CPG4	Esmoriz - Espinho	72
	Espinho – São Félix	38
	Francelos - Valadares	28
	Maceda - Cortegaça	48
	Madalena - Coimbrões	16
	Miramar - Granja	48
	São Félix – Granja	40
	Valadares - Madalena	21
	<hr/>	
	Total	1045
	<hr/>	

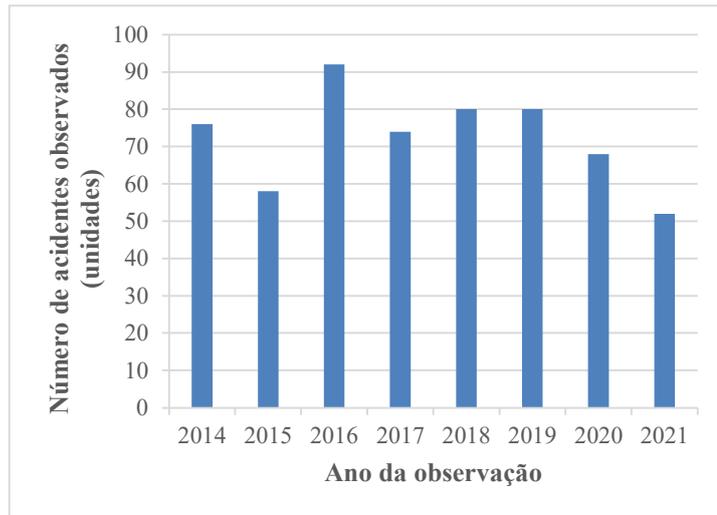
Fonte: Autor (2024)

## 2.6 ETAPA 6 – Levantamento dos dados dos acidentes nas autoestradas

Nesta etapa do estudo consiste no levantamento das informações fornecidas pela Autoridade Nacional de Segurança Rodoviária – ANSR, entre os anos de 2014 e 2021, através do Boletim Estatístico de Acidentes de Viação – BEAV que contém os índices de acidentes com o qual o estudo toma como base.

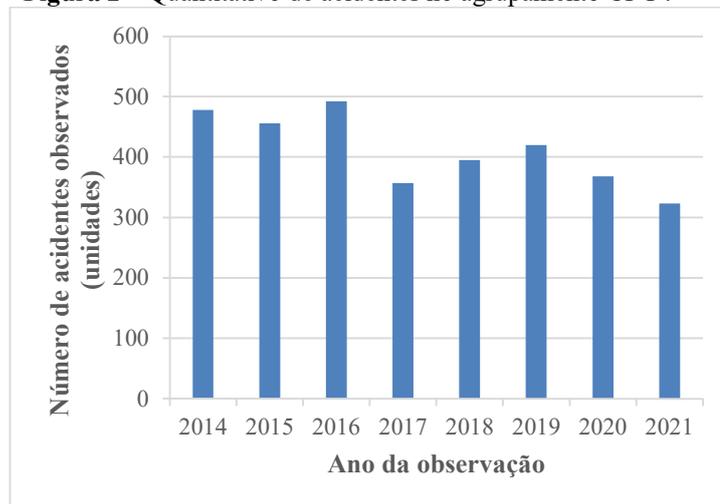
Na contabilização dos acidentes para a análise, foi considerado que os acidentes ocorrem a 500 metros à frente ou atrás do trecho a ser considerado. Nas Figuras 01 e Figuras 02 são apresentados os índices de acidentes, dentro dos agrupamentos selecionados, observados nos períodos dos anos entre 2014 e 2021 nos agrupamentos CPG3 e CPG4 respectivamente que compõe aos dados utilizados no estudo.

**Figura 1** – Quantitativo de acidentes no agrupamento CPG3



Fonte: Autor (2024)

**Figura 2** – Quantitativo de acidentes no agrupamento CPG4



Fonte: Autor (2024)

## 2.7 ETAPA 7 - Levantamento dos dados dos índices das Autoestradas

Para o modelo de previsão, se fez necessário definir as variáveis que serão analisadas a fim de possibilitar um modelo coeso com o objeto estudado. Para isso tomou-se como base os índices utilizados por Souza (2022), que contempla as características referentes ao estado do pavimento, que são indicadores do estado de conservação dos pavimentos, características geométricas em planta, que é composto por curvas circulares de raio constante, curvas de transição de raio variável e alinhamentos retos de raio infinito, como também as características geométricas em perfil, constituído por traineis e concordâncias.

Além dos fatores citados anteriormente, é de suma importância a consideração do tráfego

médio diário anual, contemplando os veículos ligeiros e pesados, como também as variáveis complementares, que engloba o número de vias e sentido dos trechos. As variáveis consideradas na base de dados são apresentadas na Tabela 5.

**Tabela 5** – Índices das autoestradas analisadas no estudo

Variável	Nomenclatura	Unidade	Média	D.P	Min	Máx
Índice Internacional de Rugosidade calculado através de uma média aritmética	IRI (Média)	m/km	1,411	0,305	0,895	3,024
Índice Internacional de Rugosidade calculado através de um valor máximo	IRI (Máximo)	m/km	1,869	0,515	0,945	3,850
Rodeiras calculado através de uma média aritmética	RD (Média)	mm	1,623	0,866	0,224	6,418
Rodeiras calculado através do valor máximo	RD (Máximo)	mm	3,359	1,806	0,550	18,850
Profundidade Média de Perfil calculado através de uma média aritmética	MPD (Média)	mm	1,696	0,295	1,038	2,554
Profundidade Média de Perfil calculado através do valor máximo	MPD (Máximo)	mm	2,083	0,428	1,250	3,750
Porcentagem da Extensão e Alinhamentos Retos no trecho de 500m	C1_AR_%	m	25,895	26,599	0,000	100,000
Porcentagem da Extensão em Curva Circular no trecho de 500m	C1_CC_%	m	30,294	21,434	0,000	100,000
Porcentagem da Extensão em Clotóide no trecho de 500m	C1_CL_%	m	31,952	20,672	0,000	76,000
Média do Inverso do Raio	R	m-1	0,003	0,002	0,000	0,007
Média do Parâmetro A	A	m	212,629	102,869	0,000	387,000
Porcentagem da Extensão em Curva Concava no trecho de 500m	C2_CCV_%	m	20,057	25,488	0,000	96,000
Porcentagem da Extensão em Curva Convexa no trecho de 500m	C2_CCX_%	m	23,759	26,681	0,000	100,000
Porcentagem da Extensão em Trainel no trecho de 500m	C2_T_%	m	44,325	30,613	0,000	100,000
Média do Raio da Curva	RC	m	988,527	3120,984	0,000	14000,000
Coefficiente de Atrito calculado através de uma média aritmética	Atrito (Média)	mm	0,455	0,059	0,267	0,641

Coefficiente de Atrito calculado através do valor mínimo	Atrito (Mínimo)	mm	0,391	0,066	0,225	0,555
Média do número de vias	N_Vias	“dummy”	2,768	0,740	1,000	4,580
Via de desaceleração	VDC	“dummy”	0,408	0,492	0,000	1,000
Via de aceleração	VAC	“dummy”	0,399	0,490	0,000	1,000
Via de aceleração direita	VACD	“dummy”	0,002	0,047	0,000	1,000
Tráfego Médio Diário Anual de veículos ligeiros	TMDAlig	Unidade/ano	17176,179	7429,630	3922,468	45650,266
Tráfego Médio Diário Anual de veículos pesados	TMDApes	Unidade/ano	1349,806	491,102	597,387	2122,798
Tráfego Médio Diário Anual total	TMDAt	Unidade/ano	18525,984	7701,973	5065,266	47743,017
Número de acidentes	Acd	Un. * dia /ano	5,581	5,378	0,000	48,000

## 2.8 ETAPA 8 - Elaboração da base de dados do estudo

Após a coleta dos dados foi feito o agrupamento dos índices das autoestradas, dados de localização e número de acidentes dos trechos em estudo.

Para a modelagem foi feita uma base de dados considerando as características e índices da via, tráfego e índice de incidência de acidentes, englobando todo o período de 2014 a 2021. Relativamente os índices e variáveis correspondentes as características das vias vão fazer parte de diferentes combinações de grupos. Esses grupos são definidos em função das áreas de influência de 500 metros e índices correspondentes a estas áreas.

Para a definição dos grupos, é levado em consideração que grande parte das variáveis estudadas na segurança rodoviária são indicadores de qualidades, características e não numericamente. Sendo assim, podem-se agrupar em categorias preestabelecidas resultando em análises categorizadas.

O primeiro grupo definido como G1, é composto por variáveis explicativas do estado do pavimento. O segundo grupo denominado G2 compõe todas as características geométricas do trecho. O terceiro grupo, G3, é o conjunto de variáveis complementares e dados relacionado ao tráfego de veículos, como pode ser observado na Tabela 6.

**Tabela 6** – Classificação das variáveis utilizadas

Grupo	Classificação	Variável
G1		IRI (Média)

		IRI (Máximo)
		RD (Média)
		RD (Máximo)
	ESTADO DO PAVIMENTO	MPD (Média)
		MPD (Máximo)
		Atrito (Média)
		Atrito (Mínimo)
		C1_AR_%
		C1_CC_%
		C1_CL_%
		R
G2		A
	CARACTERÍSTICAS GEOMÉTRICAS EM PLANTA	C2_CCX_%
		C2_CC_%
		C2_T_%
		RC
		N_Vias
		VDC
	VARIÁVEIS COMPLEMENTARES	VAC
		VACD
G3		Acd
		TMDA <sub>pes</sub>
	TRÁFEGO MÉDIO DIÁRIO ANUAL	TMDA <sub>t</sub>
		TMDA <sub>lig</sub>

## 2.9 ETAPA 9 – Ajuste do modelo

Para a modelagem, foi utilizado o programa R em sua versão 4.4.0 para ajustar à modelagem dos Modelos Mistos Generalizados. Iniciando com a utilização de todas as variáveis descritas anteriormente e prosseguindo com a retirada da variável menos influente, sendo retirada uma de cada vez, tendo em vista que a retirada de duas ou mais variáveis de forma única pode resultar no desequilíbrio dos cálculos.

A variável de definição de melhor modelo para a base de dados utilizada será o elemento estatístico raiz do erro quadrático médio. Este método é amplamente utilizado dentre os estudos na área de modelagem de dados para previsão de acidentes, como visto em Sampaio (2018).

Para a atual base de dados, a variável de definição é o indicativo do quão preciso o modelo está em relação com os sinistros observados.

$$REQM = \sqrt{\frac{\sum_{i=1}^n (y_i - \rho_i)^2}{n}}$$

Onde:

REQM = Raiz do erro quadrático médio;

$y_i$  = Variável resposta estimada para a interseção  $i$ ;

$\rho_i$  = Variável resposta observada para a interseção  $i$ ;

$n$  = Número de interseções consideradas;

$i = 1, 2, \dots, n$ .

## 2.10 Definição das variáveis da modelagem

As variáveis utilizadas na modelagem contêm informações sobre a geometria, tráfego e características das autoestradas escolhidas. Essas variáveis foram calculadas sob influência dos trechos de 500 metros definidos.

A seguir, é apresentada a definição das variáveis escolhidas para o modelo. Estas definições também são encontradas na modelagem feita por Souza (2022).

### 2.10.1 IRI – Índice Internacional de Rugosidade

O IRI indica o quão irregular é a via, o valor é medido a cada 100 metros. Para a base de dados o valor do índice é calculado a média aritmética dos valores de cada seguimento.

Para este índice foi elaborado duas variáveis, IRI (Média) que é a média aritmética do IRI a cada 500m, e IRI (Máximo) que é o valor máximo em cada agrupamento de 500m.

### 2.10.2 RD - Rodeiras

Para considerar a irregularidade transversal, é utilizado o indicador de profundidade das rodeiras. Sendo as medições feitas a cada 10 metros, o valor utilizado para a modelagem é uma média dessas medições no trecho de 500m.

Para este índice foi elaborado duas variáveis, RD (Média) que é a média aritmética do IRI a cada 500m, e RD (Máximo) que é o valor máximo em cada agrupamento de 500m.

### 2.10.3 MPD – Profundidade Média de Perfil

Sendo uma medida de rugosidade, este indicador caracteriza a textura da superfície. Sendo

determinados a cada 10 metros, foi calculado a média aritmética para o agrupamento de 500 metros.

Para este índice foi elaborado duas variáveis, MPD (Média) que é a média aritmética do IRI a cada 500m, e MPD (Mínimo) que é o valor mínimo em cada agrupamento de 500m.

#### **2.10.4 AR – Extensão de Alinhamentos Retos**

Definido como elementos de raio infinito que permitem movimento uniformes e facilitam a ultrapassagens entre veículos, porém dificultam a análise sobre velocidades e distâncias em relação a outros veículos. Para esta variável foi considerada a porcentagem de sua presença em relação ao CL e ao CC.

#### **2.10.5 CL – Extensão da Clotóide**

Sendo curvas de transição que possibilitam a ligação entre alinhamentos retos e curva circulares, as Clotóides garantem uma variação gradual do raio. Sua implementação permite a elevação do nível de segurança da via, além de permitir uma maior visualização do traçado. Para esta variável foi considerada a porcentagem de sua presença em relação ao AR e ao CC.

#### **2.10.5 CC – Extensão da Curva Circular**

Definido como elementos que permitem efetuar a concordância entre alinhamentos retos e consecutivos, facilitando a visibilidade e percepção da via ao condutor. As curvas circulares são definidas pelo seu raio. Para esta variável foi considerada a porcentagem de sua presença em relação ao CL e ao AR.

#### **2.10.6 R - Média do Inverso do Raio**

Por conta do alto valor do raio das curvas das autoestradas, para esta variável é calculado o inverso do raio a fim de evitar valores infinitos. Para esta variável foi considerada a média do trecho de 500 metros.

#### **2.10.7 A – Média do Parâmetro A**

O parâmetro A é um indicativo da qualidade da percepção da autoestrada na visão do condutor, e conseqüentemente também define a Clotóide. Para esta variável foi considerada a média do trecho de 500 metros.

#### **2.10.8 T – Extensão em Trainel**

Ao analisar a rasante, o trainel são os elementos mais simples onde sua inclinação é constante. Para esta variável foi considerada a porcentagem de sua presença em relação ao CCX e ao CCV.

#### **2.10.9 CCX – Extensão em Curva Convexa**

Utilizadas para assegurar uma distância de visibilidade adequada como forma para diminuir o índice de acidentes. Para esta variável foi considerada a porcentagem de sua presença em relação ao T e ao CCV.

#### **2.10.10 CCV – Extensão em Curva Concava**

Utilizadas para assegurar uma visibilidade noturna adequada na via e comodidade. Para esta variável foi considerada a porcentagem de sua presença em relação ao T e ao CCX.

#### **2.10.11 RC – Raio da Curva**

Índice correspondente ao raio da curva das vias. Para esta variável foi utilizado a média do raio no agrupamento de 500m.

#### **2.10.12 TMDA – Tráfego Médio Diário Anual**

É o valor do tráfego anual que há na autoestrada, separado pelo número de dias do ano. Este valor pode variar de acordo com o ano em que foi medido.

Para a base de dados foi feita a separação em três categorias, TMDA<sub>lig</sub> para o tráfego de veículos ligeiros, TMDA<sub>pes</sub> para o tráfego de veículos pesados, e TMDA<sub>at</sub> para o tráfego médio

diário anual total.

### 2.10.13 Variáveis complementares

São as variáveis utilizadas para auxiliar a modelagem dos dados, não são índices técnicos, mas contém importâncias geométricas e facilitam a localização dos trechos. Estas variáveis foram organizadas na tabela na Tabela 7:

**Tabela 7 – Variáveis complementares**

<b>Variável</b>	<b>Definição</b>
N_Vias	Número de Vias
VDC	Via Central Direita
VAC	Via de Aceleração
VACD	Via de Aceleração Direita

---

## 3 RESULTADOS E DISCUSSÕES

Na modelagem de previsão de acidentes ocorridos nas autoestradas estudadas foram utilizadas as equações de modelos lineares generalizados mistos com as distribuições de Poisson e Binomial negativa. Os dados foram recolhidos através dos diários da Autoridade Nacional de Segurança Rodoviária entre os anos de 2014 e 2021. Então, para este capítulo é apresentado os modelos de previsão de acidades para os trechos citados.

### 3.1 Modelos para trechos de 500 metros

Nesta seção são apresentados os modelos finais de previsão de acidentes ajustados para os trechos de 500 metros. Para a modelagem foi considerado somente a variáveis Acd como variável dependente, e também as variáveis explicativas, definidas em função da geometria e características dos trechos utilizados. Na Tabela 8 são representadas as combinações para os modelos.

**Tabela 8 – Combinação dos modelos de previsão**

<b>Nº</b>	<b>Função</b>	<b>Variável dependente</b>
1	Poisson	Acd
2	Binomial Negativa	Acd

Para as combinações foram utilizadas a mesma variável dependente, porém alternando a função de distribuição e o agrupamento de dados. Em cada uma das combinações o modelo final foi ajustado para que todas as variáveis tenham um nível de significância de 5%. Para este ajuste foi calculado inicialmente o modelo com todas as 25 variáveis do estudo, e em seguida foi realizada a remoção da variável menos significativa a cada cálculo do modelo. A seguir são apresentados na Tabela 9 os modelos para as duas combinações propostas no estudo.

**Tabela 9** – Modelo de previsão de acidentes para a Combinação 1

Nº	Variável dependente	Parâmetros	Estimativa	Desvio Padrão	P
1	Acd	Intercepto	1,215	0,123	< 2e-16
		C1_AR_%	0,487	0,043	< 2e-16
		C1_CC_%	0,413	0,036	< 2e-16
		C1_CL_%	0,441	0,033	< 2e-16
		R	-0,128	0,031	5,45e-05
		A	-0,240	0,042	1,11e-08
		Atrito (Média)	0,077	0,034	0,022
		Atrito (Máximo)	-0,108	0,035	0,002
		TMDAt	0,218	0,031	5,14e-12
Número de observações: 1046					

No modelo de Combinação 1, verificou-se que as variáveis C1\_AR\_%, C1\_CC\_%, C1\_CL\_% e Atrito (Média) contribuem positivamente para a ocorrência de acidentes. E as variáveis R, A e Atrito (Máximo) contribuem negativamente para as ocorrências observadas, conforme exibido na Tabela 10.

**Tabela 10** – Modelo de previsão de acidentes para a Combinação 2

Nº	Variável dependente	Parâmetros	Estimativa	Desvio Padrão	P
2	Acd	Intercepto	1,196	0,120	< 2e-16
		C1_AR_%	0,465	0,057	3,53e-16
		C1_CC_%	0,385	0,048	2,19e-15
		C1_CL_%	0,457	0,044	< 2e-16
		R	-0,105	0,042	0,0125
		A	-0,235	0,052	7,90e-06
		TMDAt	0,278	0,050	4,36e-08
		Número de observações: 1046			

No modelo de Combinação 2, verificou-se que as variáveis C1\_AR\_%, C1\_CC\_%, e C1\_CL\_% contribuem positivamente para a ocorrência de acidentes. E as variáveis R, e A contribuem negativamente para as ocorrências observadas.

### 3.2 Análise dos resultados dos modelos para trechos de 500 metros

Na Tabela 11, é descrito o resumo dos resultados dos modelos de previsão de acidentes desenvolvidos e ajustados para trechos de 500 metros, sendo apresentados a função utilizada, a variável dependente, o tamanho da amostra e a variável indicativa de precisão do modelo, erro quadrático médio.

**Tabela 11** – Elementos estatísticos dos modelos finais ajustados

<b>Combinação</b>	<b>Função</b>	<b>Tamanho da amostra</b>	<b>REQM</b>
1	Poisson	1046	3,023
2	Binomial Negativa	1046	2,993

O raio quadrático médio foi o método escolhido para definir o modelo mais eficiente para o estudo, já que esta variável pode ser aplicada para validar a precisão do modelo. Ou seja, no estudo em que ela foi aplicada, quanto mais próximo de 0 o seu valor estiver, mais preciso o modelo.

Logo, ao analisar a Tabela 11, pode-se conferir que ambos os REQM estão próximos do valor 3. A Combinação 01 apresentou um índice de 3,023 e a Combinação 2 um índice de 2,993. Portanto, para o estudo em que foi aplicado, a modelagem utilizando a função binomial negativa é a mais eficaz para a previsão de acidentes viários.

---

## 4 CONCLUSÕES

Este trabalho apresenta uma análise dos fatores contribuintes na ocorrência de acidentes nas autoestradas da região norte de Portugal, mais precisamente nas concessões da Costa de Prata Grupo 3 e Grupo 4.

A identificação das variáveis que contribuem positivamente e negativamente na ocorrência dos sinistros em segmentos de trechos de 500 metros antes e após o ponto de ocorrência foi o foco do presente estudo. Colaborando para viabilizar políticas e metodologias de prevenção e

controle de acidentes em regiões de maiores periculosidades. Para tanto, modelos lineares generalizados mistos utilizando as funções de Poisson e Binomial Negativa foram aplicados nos dados levantados por este estudo, sendo variáveis indicativas do estado do pavimento, índices das características geométricas da via, e tráfego médio diário anual.

Para as modelagens finais dos modelos desenvolvidos, foi evidenciado que as variáveis relacionadas ao raio da curva, parâmetro A, alinhamentos retos, extensão da curva circular e Clotóide são parâmetros significativos para a previsão de acidentes em ambos os modelos.

Com a aplicação do erro quadrático médio, definiu-se que a modelagem mais precisa para os dados do estudo é a que utiliza a função binomial negativa, apresentando um erro de 2,993. Porém ao comparado com a modelagem que utilizou a função de Poisson, a diferença entre os erros foi somente de 0,03, revelando que embora com um erro maior a utilização da função de Poisson ainda resultaria em resultados próximos dos observados.

---

## REFERÊNCIAS

Antunes, M., Marecos, V. **Indicadores de Desempenho para Pavimentos Rodoviários.**V Congresso Rodoviario Portugues, Centro Rodoviario Portugues, Estoril, 2008.

AUTORIDADE NACIONAL DE SEGURANÇA RODOVIÁRIA. **Sinistralidade rodoviária.** ANSR, 2021.

RAFAEL ANDRÉ BARRETO DA SILVA. **Modelos de Previsão de Acidentes para Interseções de Três Ramos em Zonas Urbanas e Periurbanas.** Dissertação (Mestrado em Engenharia Urbana) – Universidade do Minho, 2015.

MANNERING, F.L., BHAT, C.R., **Analytic methods in accident research: Methodological frontier and future directions.** Analytic Methods in Accident Research. Vol 1, 2014.

ZAIHRA, T. **Modeling the Number of Research Papers Produced by Graduate Students Using Zero-Inflated Models.** Case Studies In Business, Industry And Government Statistics. CS-BIGS, 2012.

SAMPEDRO, A. **Procedimento para avaliação e análise da segurança de tráfego em vias expressas urbanas.** Tese (Doutorado em Engenharia de Transportes). UFRJ, 2010.

MAIA, J. **Uma análise sistêmica dos acidentes de trânsito no Brasil.** Dissertação (Mestrado em Engenharia Civil e Ambiental). João Pessoa: Universidade Federal da Paraíba, 1995.

MÁRIO TALAIA & NUNO AMARO. **A segurança rodoviária e os acidentes – estudo de caso.** Territorium 19, 2012, 209-214.

NODARI, C. T., & LINDAU, L. A. **Auditorias de segurança viária.** In: Transportes. vol. 9, n. 2, 2001. P 48-66.

JOCILENE OTÍLIA DA COSTA. **Desenvolvimento dum modelo de previsão de acidentes.** Tese (Doutorado em Engenharia Civil) – Universidade do Minho, 2018.

EDUARDO SOUZA SAMPAIO. **Modelos de previsão da severidade dos acidentes viários em rodovias nacionais de Portugal.** Dissertação (Graduação em Engenharia Civil) – Universidade Federal de Pernambuco, 2018.

CÁTIA MARIA OLIVEIRA DE SOUZA. **Modelos de Regressão para Dados de Contagem: Aplicação a Dados de Contagem Rodoviários.** Dissertação (Mestrado em Estatística) – Universidade do Minho, 2022.

LORD, D., & MANNERING, F.. **The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives.** Transportation Research Part A, 44, 291-305, 2010.

MÁRIO TALAIA, **A segurança rodoviária e os acidentes – estudo de caso.** Territorium 19, 209-2014, 2012.

LUCAS HENRIQUE DA CRUZ FERREIRA

**ANÁLISE DOS FATORES CONTRIBUINTES NA OCORRÊNCIA DE ACIDENTES  
EM AUTOESTRADAS DA REGIÃO NORTE DE PORTUGAL**

Trabalho de Conclusão de Curso apresentado à  
Coordenação do Curso de Engenharia Civil do  
Campus Agreste da Universidade Federal de  
Pernambuco – UFPE, na modalidade de artigo  
científico, como requisito parcial para obtenção  
do grau de Bacharel em Engenharia Civil.

**Área de concentração:** Transportes

Aprovado em 17 de maio de 2024.

**BANCA EXAMINADORA**

---

Profa. Dra. Jocilene Otília da Costa (Orientadora)  
Universidade Federal de Pernambuco

---

Profa. Dra. Dannúbia Ribeiro Pires (Avaliadora)  
Universidade Federal de Pernambuco

---

Prof. Dr. Edevaldo Miguel Alves (Avaliador)  
Universidade Federal de Pernambuco